```
Qy          1 MRPLLVLLLLGLAAGSPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MRPLLVLLLLGLAAGSPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPG 60

Qy         61 APGEKGEGGRPGLPGPRGDPGPRGEAGPAGPTGPAGECSVPPRSAFSAKRSESRVPPPSD 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 APGEKGEGGRPGLPGPRGDPGPRGEAGPAGPTGPAGECSVPPRSAFSAKRSESRVPPPSD 120

Qy        121 APLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGESIASFFQ 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 APLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGESIASFFQ 180

Qy        181 FFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWHSSP 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 FFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWHSSP 240

Qy        241 VFA 243
              |||
Db        241 VFA 243
```

RESULT 2
APM1_MOUSE
ID   APM1_MOUSE     STANDARD;      PRT;   247 AA.
AC   Q60994; Q62400; Q9DC68;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Adiponectin precursor (30 kDa adipocyte complement-related protein)
DE   (ACRP30) (Adipocyte specific protein AdipoQ).
GN   APM1 OR ACRP30 OR ADIPOQ.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Adipocyte;
RX   MEDLINE=96070757; PubMed=7592907;
RA   Scherer P.E., Williams S., Fogliano M., Baldini G., Lodish H.F.;
RT   "A novel serum protein similar to C1q, produced exclusively in
RT   adipocytes.";
RL   J. Biol. Chem. 270:26746-26749(1995).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Fibroblast;
RX   MEDLINE=96209999; PubMed=8631877;
RA   Hu E., Liang P., Spiegelman B.M.;
RT   "AdipoQ is a novel adipose-specific gene dysregulated in obesity.";
RL   J. Biol. Chem. 271:10697-10703(1996).
RN   [3]
RP   SEQUENCE FROM N.A.
RX   PubMed=11162643;
RA   Das K., Lin Y., Widen E., Zhang Y., Scherer P.E.;
RT   "Chromosomal localization, expression pattern, and promoter analysis
RT   of the mouse gene encoding adipocyte-specific secretory protein
RT   Acrp30.";
RL   Biochem. Biophys. Res. Commun. 280:1120-1129(2001).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Heart;
RX   MEDLINE=21085660; PubMed=11217851;
RA   Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA   Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA   Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA   Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA   Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA   Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA   Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA   Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA   Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA   Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA   Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA   Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA   Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA   Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA   Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA   Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA   Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA   Hayashizaki Y.;
RT   "Functional annotation of a full-length mouse cDNA collection.";
RL   Nature 409:685-690(2001).
RN   [5]
RP   FUNCTION.
RX   MEDLINE=21372498; PubMed=11479627;
RA   Yamauchi T., Kamon J., Waki H., Terauchi Y., Kubota N., Hara K.,
RA   Mori Y., Ide T., Murakami K., Tsuboyama-Kasaoka N., Ezaki O.,
RA   Akanuma Y., Gavrilova O., Vinson C., Reitman M.L., Kagechika H.,
RA   Shudo K., Yoda M., Nakano Y., Tobe K., Nagai R., Kimura S., Tomita M.,
RA   Froguel P., Kadowaki T.;
RT   "The fat-derived hormone adiponectin reverses insulin resistance
RT   associated with both lipoatrophy and obesity.";
RL   Nat. Med. 7:941-946(2001).
RN   [6]
RP   FUNCTION.
RX   MEDLINE=21372499; PubMed=11479628;
RA   Berg A.H., Combs T.P., Du X., Brownlee M., Scherer P.E.;
RT   "The adipocyte-secreted protein Acrp30 enhances hepatic insulin
RT   action.";
RL   Nat. Med. 7:947-953(2001).
CC   -!- FUNCTION: IMPORTANT NEGATIVE REGULATOR IN HEMATOPOIESIS AND IMMUNE
CC       SYSTEMS; MAY BE INVOLVED IN ENDING INFLAMMATORY RESPONSES THROUGH
CC       ITS INHIBITORY FUNCTIONS. INHIBITS ENDOTHELIAL NF-KAPPAB SIGNALING
CC       THROUGH A CAMP-DEPENDENT PATHWAY. INHIBITS TNF-ALPHA-INDUCED
CC       EXPRESSION OF ENDOTHELIAL ADHESION MOLECULES. INVOLVED IN THE
CC       CONTROL OF FAT METABOLISM AND INSULIN SENSITIVITY.
CC   -!- SUBUNIT: HOMOOLIGOMER.
CC   -!- SUBCELLULAR LOCATION: SECRETED IN PLASMA.
CC   -!- TISSUE SPECIFICITY: SYNTHESIZED EXCLUSIVELY BY ADIPOCYTES AND
CC       SECRETED INTO SERUM.
CC   -!- INDUCTION: DURING HORMONE-INDUCED ADIPOSE DIFFERENTIATION AND
CC       ACTIVATED BY INSULIN.
CC   -!- SIMILARITY: CONTAINS 1 COLLAGENOUS DOMAIN.
CC   -!- SIMILARITY: CONTAINS 1 C1Q DOMAIN.
CC   ---------------------------------------------------------------------

CC   ---------------------------------------------------------------------
DR   EMBL; U37222; AAA80543.1; -.
DR   EMBL; U49915; AAB06706.1; -.
DR   EMBL; AF304466; AAK13417.1; -.
DR   EMBL; AK003138; BAB22597.1; -.
DR   MGD; MGI:106675; Acrp30.
DR   InterPro; IPR001073; C1q.
DR   InterPro; IPR000087; Collagen.
DR   Pfam; PF00386; C1q; 1.
DR   Pfam; PF01391; Collagen; 1.
DR   PRINTS; PR00007; COMPLEMNTC1Q.
DR   SMART; SM00110; C1Q; 1.
DR   PROSITE; PS01113; C1Q; 1.
KW   Hormone; Collagen; Signal; Repeat; Hydroxylation; Plasma;
KW   Polymorphism.
FT   SIGNAL       1     17       POTENTIAL.
FT   CHAIN       18    247       ADIPONECTIN.
FT   DOMAIN      45    110       COLLAGEN-LIKE.
FT   DOMAIN     111    247       C1Q.
FT   DISULFID    39     39       INTERCHAIN (BY SIMILARITY).
FT   MOD_RES     47     47       HYDROXYLATION (BY SIMILARITY).
FT   MOD_RES     50     50       HYDROXYLATION (BY SIMILARITY).
FT   MOD_RES     56     56       HYDROXYLATION (BY SIMILARITY).

OM protein - protein search, using sw model

Run on: June 18, 2003, 14:46:42 ; Search time 22 Seconds
(without alignments)
458.125 Million cell updates/sec

Title: US-09-943-851A-42
Perfect score: 1325
Sequence: 1 MRPLLVLLLLGLAAGSPPLD.....DSTFSGFLVYSDWHSSPVFA 243

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 1325 | 100.0 | 243 | 1 | CQT5_HUMAN | Q9bxj0 homo sapien |
| 2 | 433.5 | 32.7 | 247 | 1 | APM1_MOUSE | Q60994 mus musculu |
| 3 | 424 | 32.0 | 244 | 1 | APM1_HUMAN | Q15848 homo sapien |
| 4 | 418.5 | 31.6 | 680 | 1 | CA1A_HUMAN | Q03692 homo sapien |
| 5 | 417 | 31.5 | 674 | 1 | CA1A_CHICK | P08125 gallus gall |
| 6 | 412 | 31.1 | 246 | 1 | C1QC_MOUSE | Q02105 mus musculu |
| 7 | 411.5 | 31.1 | 674 | 1 | CA1A_BOVIN | P23206 bos taurus |
| 8 | 410.5 | 31.0 | 289 | 1 | CQT7_HUMAN | Q9bxj2 homo sapien |
| 9 | 408 | 30.8 | 680 | 1 | CA1A_MOUSE | Q05306 mus musculu |
| 10 | 403 | 30.4 | 285 | 1 | CQT2_HUMAN | Q9bxj5 homo sapien |
| 11 | 397 | 30.0 | 245 | 1 | C1QC_HUMAN | P02747 homo sapien |
| 12 | 374 | 28.2 | 744 | 1 | CA18_RABIT | P14282 oryctolagus |
| 13 | 373 | 28.2 | 419 | 1 | COLE_LEPMA | P98085 lepomis mac |
| 14 | 368 | 27.8 | 744 | 1 | CA18_HUMAN | P27658 homo sapien |
| 15 | 366.5 | 27.7 | 743 | 1 | CA18_MOUSE | Q00780 mus musculu |
| 16 | 363 | 27.4 | 635 | 1 | CA28_HUMAN | P25067 homo sapien |
| 17 | 362 | 27.3 | 251 | 1 | C1QB_HUMAN | P02746 homo sapien |
| 18 | 360.5 | 27.2 | 253 | 1 | C1QB_RAT | P31721 rattus norv |
| 19 | 350 | 26.4 | 253 | 1 | C1QB_MOUSE | P14106 mus musculu |
| 20 | 314 | 23.7 | 245 | 1 | C1QA_HUMAN | P02745 homo sapien |
| 21 | 314 | 23.7 | 245 | 1 | C1QA_MOUSE | P98086 mus musculu |
| 22 | 298.5 | 22.5 | 255 | 1 | GLIC_MOUSE | Q9esn4 mus musculu |
| 23 | 285.5 | 21.5 | 258 | 1 | C1RF_HUMAN | O75973 homo sapien |
| 24 | 282 | 21.3 | 258 | 1 | C1RF_MOUSE | O88992 mus musculu |
| 25 | 280.5 | 21.2 | 215 | 1 | HP25_TAMSI | Q06576 tamias sibi |
| 26 | 277.5 | 20.9 | 215 | 1 | HP27_TAMSI | Q06577 tamias sibi |
| 27 | 264.5 | 20.0 | 246 | 1 | CQT3_HUMAN | Q9bxj4 homo sapien |
| 28 | 261.5 | 19.7 | 196 | 1 | HP20_TAMSI | Q06575 tamias sibi |
| 29 | 230.5 | 17.4 | 1049 | 1 | CA13_BOVIN | P04258 bos taurus |
| 30 | 229.5 | 17.3 | 281 | 1 | CQT1_HUMAN | Q9bxj1 homo sapien |
| 31 | 228 | 17.2 | 684 | 1 | CA39_HUMAN | Q14050 homo sapien |
| 32 | 224 | 16.9 | 636 | 1 | CA13_RAT | P13941 rattus norv |
| 33 | 223 | 16.8 | 1464 | 1 | CA13_MOUSE | P08121 mus musculu |
| 34 | 222.5 | 16.8 | 1758 | 1 | CA24_CAEEL | P17140 caenorhabdi |
| 35 | 221.5 | 16.7 | 170 | 1 | CA28_MOUSE | P25318 mus musculu |
| 36 | 221 | 16.7 | 1019 | 1 | CA16_CHICK | P20785 gallus gall |
| 37 | 221 | 16.7 | 1262 | 1 | CA13_CHICK | P12105 gallus gall |
| 38 | 220 | 16.6 | 248 | 1 | PSPA_CANFA | P06908 canis famil |
| 39 | 220 | 16.6 | 360 | 1 | CCD2_CAEEL | P35799 caenorhabdi |
| 40 | 219.5 | 16.6 | 247 | 1 | PSPA_RABIT | P12842 oryctolagus |
| 41 | 219.5 | 16.6 | 1466 | 1 | CA13_HUMAN | P02461 homo sapien |
| 42 | 219 | 16.5 | 671 | 1 | CA11_RAT | P02454 rattus norv |
| 43 | 218.5 | 16.5 | 1453 | 1 | CA11_MOUSE | P11087 mus musculu |
| 44 | 217.5 | 16.4 | 1516 | 1 | CA1H_HUMAN | P39060 homo sapien |
| 45 | 217 | 16.4 | 747 | 1 | CA12_BOVIN | P02459 bos taurus |

ALIGNMENTS

RESULT 1
CQT5_HUMAN
ID CQT5_HUMAN STANDARD; PRT; 243 AA.
AC Q9BXJ0; Q9UFX4;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Complement-c1q tumor necrosis factor-related protein 5 precursor.
GN C1QTNF5 OR CTRP5.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Sheppard P.O., Humes J.M.;
RT "Homo sapiens complement-c1q tumor necrosis factor-related protein.";
RL Submitted (DEC-2000) to the EMBL/GenBank/DDBJ databases.
RN [2]
RP SEQUENCE OF 25-243 FROM N.A.
RC TISSUE=Uterus;
RA Ottenwaelder B., Obermaier B., Mewes H.-W., Gassenhuber J.,
RA Wiemann S.;
RL Submitted (AUG-1999) to the EMBL/GenBank/DDBJ databases.
CC -!- SIMILARITY: CONTAINS 1 COLLAGENOUS DOMAIN.
CC -!- SIMILARITY: CONTAINS 1 C1Q DOMAIN.
CC ---
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC ---
DR EMBL; AF329841; AAK17965.1; -.
DR EMBL; AL110261; CAB53702.1; -.
DR Genew; HGNC:14344; C1QTNF5.
DR InterPro; IPR001073; C1q.
DR InterPro; IPR000087; Collagen.
DR Pfam; PF00386; C1q; 1.
DR Pfam; PF01391; Collagen; 1.
DR PRINTS; PR00007; COMPLEMNTC1Q.
DR SMART; SM00110; C1Q; 1.
DR PROSITE; PS01113; C1Q; FALSE_NEG.
KW Collagen; Signal.
FT SIGNAL 1 15 POTENTIAL.
FT CHAIN 16 243 COMPLEMENT-C1Q TUMOR NECROSIS FACTOR-
FT RELATED PROTEIN 5.
FT DOMAIN 30 95 COLLAGEN-LIKE.
FT DOMAIN 97 243 C1Q.
SQ SEQUENCE 243 AA; 25298 MW; 7CCDA65CDA7EB784 CRC64;

Query Match 100.0%; Score 1325; DB 1; Length 243;
Best Local Similarity 100.0%; Pred. No. 1.7e-85;
Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
FT   MOD_RES      65     65       HYDROXYLATION (BY SIMILARITY).
FT   MOD_RES      79     79       HYDROXYLATION (BY SIMILARITY).
FT   MOD_RES      98     98       HYDROXYLATION (BY SIMILARITY).
FT   MOD_RES     107    107       HYDROXYLATION (BY SIMILARITY).
FT   VARIANT     113    113       M -> V.
FT   CONFLICT     50     50       P -> S (IN REF. 2).
FT   CONFLICT     74     74       A -> S (IN REF. 2).
FT   CONFLICT    117    117       A -> G (IN REF. 2).
FT   CONFLICT    148    148       G -> N (IN REF. 2).
FT   CONFLICT    243    243       Y -> F (IN REF. 2).
SQ   SEQUENCE   247 AA;  26841 MW;  137B687D873988C4 CRC64;

  Query Match           32.7%;  Score 433.5;  DB 1;  Length 247;
  Best Local Similarity  40.6%;  Pred. No. 1.2e-23;
  Matches  102;  Conservative   32;  Mismatches   92;  Indels   25;  Gaps    8;

Qy          1 MRPLLVLLLL-GLAAGSPPLDDNKIPSLCPGHPG-----LPGTPGHHGSQGLPGRDGRDG 54
              :: || ||:|   |       :   |:| |   |     : | ||| |  | |||||||
Db          4 LQALLFLLILPSHAEDDVTTTEELAPALVPPPKGTCAGWMAGIPGHPGHNGTPGRDGRD- 62

Qy         55 RDGAPGAPGEKGEGGRPGLPGPRGD---------PGPRGEAGPAGPTGPAGECSVPPRSA 105
                | |||||| |  || ||:|:           ||||  |  |  |  || :   |||
Db         63 -----GTPGEKGEKGDAGLLGPKGETGDVGMTGAEGPRGFPGTPGRKGEPGEAAYMYRSA 117

Qy        106 FSAKRSESRVPPPSDAPLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQ 165
              ||    |:||  | : |: | ::  |:| |||  |||| | :||:|||: | |||   ::
Db        118 FSV-GLETRVTVP-NVPIRFTKIFYNQQNHYDGSTGKFYCNIPGLYYFSYHITVYMKDVK 175

Qy        166 FDLVKNGESIASFFQFFGGWPKPASLSGGAMVRLEPEDQVWVQV-GVGDYIGIYASIKTD 224
                | |  :::   :  :          ||  :: ||  ||||:|| | ||: |:||    |
Db        176 VSLFKKDKAVLFTYDQYQE-KNVDQASGSVLLHLEVGDQVWLQVYGDGDHNGLYADNVND 234

Qy        225 STFSGFLVYSD 235
              |||:|||:| |
Db        235 STFTGFLLYHD 245


RESULT 3
APM1_HUMAN
ID   APM1_HUMAN     STANDARD;      PRT;   244 AA.
AC   Q15848;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Adiponectin precursor (30 kDa adipocyte complement-related protein)
DE   (ACRP30) (Adipose most abundant gene transcript 1) (apM-1) (Gelatin-
DE   binding protein).
GN   APM1 OR ACRP30 OR GBP28.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Adipose tissue;
RX   MEDLINE=96224171; PubMed=8619847;
RA   Maeda K., Okubo K., Shimomura I., Funahashi T., Matsuzawa Y.,
RA   Matsubara K.;
RT   "cDNA cloning and expression of a novel adipose specific collagen-like
RT   factor, apM1 (AdiPose Most abundant Gene transcript 1).";
RL   Biochem. Biophys. Res. Commun. 221:286-289(1996).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99196984; PubMed=10095105;
RA   Saito K., Tobe T., Minoshima S., Asakawa S., Sumiya J., Yoda M.,
RA   Nakano Y., Shimizu N., Tomita M.;
RT   "Organization of the gene for gelatin-binding protein (GBP28).";
RL   Gene 229:67-73(1999).
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99333693; PubMed=10403784;
RA   Schaeffler A., Orso E., Palitzsch K.D., Buechler C., Drobnik W.,
RA   Fuerst A., Schoelmerich J., Schmitz G.;
RT   "The human apM-1, an adipocyte-specific gene linked to the family of
RT   TNF's and to genes expressed in activated T cells, is mapped to
RT   chromosome 1q21.3-q23, a susceptibility locus identified for familial
RT   combined hyperlipidemia (FCH).";
RL   Biochem. Biophys. Res. Commun. 260:416-425(1999).
RN   [4]
RP   CHARACTERIZATION.
RX   MEDLINE=20417747; PubMed=10961870;
RA   Yokota T., Oritani K., Takahashi I., Ishikawa J., Matsuyama A.,
RA   Ouchi N., Kihara S., Funahashi T., Tenner A.J., Tomiyama Y.,
RA   Matsuzawa Y.;
RT   "Adiponectin, a new member of the family of soluble defense collagens,
RT   negatively regulates the growth of myelomonocytic progenitors and the
RT   functions of macrophages.";
RL   Blood 96:1723-1732(2000).
RN   [5]
RP   CHARACTERIZATION.
RX   MEDLINE=20440368; PubMed=10982546;
RA   Ouchi N., Kihara S., Arita Y., Okamoto Y., Maeda K., Kuriyama H.,
RA   Hotta K., Nishida M., Takahashi M., Muraguchi M., Ohmoto Y.,
RA   Nakamura T., Yamashita S., Funahashi T., Matsuzawa Y.;
RT   "Adiponectin, an adipocyte-derived plasma protein, inhibits
RT   endothelial NF-kappaB signaling through a cAMP-dependent pathway.";
RL   Circulation 102:1296-1301(2000).
RN   [6]
RP   FUNCTION.
RX   MEDLINE=21372498; PubMed=11479627;
RA   Yamauchi T., Kamon J., Waki H., Terauchi Y., Kubota N., Hara K.,
RA   Mori Y., Ide T., Murakami K., Tsuboyama-Kasaoka N., Ezaki O.,
RA   Akanuma Y., Gavrilova O., Vinson C., Reitman M.L., Kagechika H.,
RA   Shudo K., Yoda M., Nakano Y., Tobe K., Nagai R., Kimura S., Tomita M.,
RA   Froguel P., Kadowaki T.;
RT   "The fat-derived hormone adiponectin reverses insulin resistance
RT   associated with both lipoatrophy and obesity.";
RL   Nat. Med. 7:941-946(2001).
RN   [7]
RP   VARIANT ADIPONECTIN DEFICIENCY CYS-112.
RX   MEDLINE=20378830; PubMed=10918532;
RA   Takahashi M., Arita Y., Yamagata K., Matsukawa Y., Okutomi K.,
RA   Horie M., Shimomura I., Hotta K., Kuriyama H., Kihara S., Nakamura T.,
RA   Yamashita S., Funahashi T., Matsuzawa Y.;
RT   "Genomic structure and mutations in adipose-specific gene,
RT   adiponectin.";
RL   Int. J. Obes. Relat. Metab. Disord. 24:861-868(2000).
RN   [8]
RP   VARIANTS ARG-84; MET-117; THR-164; SER-221 AND PRO-241.
RX   MEDLINE=21671103; PubMed=11812766;
RA   Hara K., Boutin P., Mori Y., Tobe K., Dina C., Yasuda K., Yamauchi T.,
RA   Otabe S., Okada T., Eto K., Kadowaki H., Hagura R., Akanuma Y.,
RA   Yazaki Y., Nagai R., Taniyama M., Matsubara K., Yoda M., Nakano Y.,
RA   Kimura S., Tomita M., Kimura S., Ito C., Froguel P., Kadowaki T.;
RT   "Genetic variation in the gene encoding adiponectin is associated with
RT   an increased risk of type 2 diabetes in the Japanese population.";
RL   Diabetes 51:536-540(2002).
CC   -!- FUNCTION: IMPORTANT NEGATIVE REGULATOR IN HEMATOPOIESIS AND IMMUNE
CC       SYSTEMS; MAY BE INVOLVED IN ENDING INFLAMMATORY RESPONSES THROUGH
CC       ITS INHIBITORY FUNCTIONS. INHIBITS ENDOTHELIAL NF-KAPPAB SIGNALING
CC       THROUGH A CAMP-DEPENDENT PATHWAY. INHIBITS TNF-ALPHA-INDUCED
CC       EXPRESSION OF ENDOTHELIAL ADHESION MOLECULES. INVOLVED IN THE
CC       CONTROL OF FAT METABOLISM AND INSULIN SENSITIVITY.
CC   -!- SUBUNIT: HOMOOLIGOMER (POTENTIAL).
CC   -!- SUBCELLULAR LOCATION: SECRETED IN PLASMA.
CC   -!- TISSUE SPECIFICITY: EXPRESSED ONLY IN THE ADIPOSE TISSUE.
CC   -!- DISEASE: Defects in APM1 are the cause of adiponectin deficiency,
CC       resulting in very low concentration of plasma adiponectin.
CC       Decreased adiponectin plasma levels are associated with obesity
CC       insulin resistance, and diabetes type 2.
CC   -!- PHARMACEUTICAL: Adiponectin might be used in the treatment of
CC       diabetes type 2 and insulin resistance.
CC   -!- SIMILARITY: CONTAINS 1 COLLAGENOUS DOMAIN.
CC   -!- SIMILARITY: CONTAINS 1 C1Q DOMAIN.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
```

```
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; D45371; BAA08227.1; -.
DR   EMBL; AB012165; BAA86716.1; -.
DR   EMBL; AB012164; BAA86716.1; JOINED.
DR   EMBL; AJ131460; CAB52413.1; -.
DR   EMBL; AJ131461; CAB52413.1; JOINED.
DR   MIM; 605441; -.
DR   InterPro; IPR001073; C1q.
DR   InterPro; IPR000087; Collagen.
DR   Pfam; PF00386; C1q; 1.
DR   Pfam; PF01391; Collagen; 1.
DR   PRINTS; PR00007; COMPLEMNTC1Q.
DR   SMART; SM00110; C1Q; 1.
DR   PROSITE; PS01113; C1Q; 1.
KW   Hormone; Collagen; Signal; Repeat; Hydroxylation; Plasma;
KW   Polymorphism; Disease mutation; Obesity; Diabetes mellitus.
FT   SIGNAL        1    14       POTENTIAL.
FT   CHAIN        15   244       ADIPONECTIN.
FT   DOMAIN       42   107       COLLAGEN-LIKE.
FT   DOMAIN      108   244       C1Q.
FT   DISULFID     36    36       INTERCHAIN (BY SIMILARITY).
FT   MOD_RES      44    44       HYDROXYLATION (BY SIMILARITY).
FT   MOD_RES      47    47       HYDROXYLATION (BY SIMILARITY).
FT   MOD_RES      53    53       HYDROXYLATION (BY SIMILARITY).
FT   MOD_RES      62    62       HYDROXYLATION (BY SIMILARITY).
FT   MOD_RES      71    71       HYDROXYLATION (BY SIMILARITY).
FT   MOD_RES      76    76       HYDROXYLATION (BY SIMILARITY).
FT   MOD_RES      86    86       HYDROXYLATION (BY SIMILARITY).
FT   MOD_RES      95    95       HYDROXYLATION (BY SIMILARITY).
FT   MOD_RES     104   104       HYDROXYLATION (BY SIMILARITY).
FT   VARIANT      84    84       G -> R.
FT                               /FTId=VAR_013273.
FT   VARIANT     112   112       R -> C (IN ADIPONECTIN DEFICIENY).
FT                               /FTId=VAR_013274.
FT   VARIANT     117   117       V -> M.
FT                               /FTId=VAR_013275.
FT   VARIANT     164   164       I -> T.
FT                               /FTId=VAR_013276.
FT   VARIANT     221   221       R -> S.
FT                               /FTId=VAR_013277.
FT   VARIANT     241   241       H -> P.
FT                               /FTId=VAR_013278.
SQ   SEQUENCE   244 AA;  26414 MW;  64D8C6C1204B1018 CRC64;

  Query Match             32.0%;  Score 424;  DB 1;  Length 244;
  Best Local Similarity   40.5%;  Pred. No. 5.3e-23;
  Matches  106;  Conservative   26;  Mismatches   72;  Indels   58;  Gaps   10;

Qy          6 VLLLLGLAAGSPPLDDNKIPSLCPGH--------PG-------------LPGTPGHHGSQ 44
              ||||| |                  |||        ||             : | ||| |
Db          7 VLLLLAL----------------PGHDQETTTQGPGVLLPLPKGACTGWMAGIPGHPGHN 50

Qy         45 GLPGRDGRDGRDGAPGAPGEKGEGGRPGLPGPRGD---------PGPRGEAGPAGPTGPA 95
              | |||||||      | |||||| | ||| ||:||          ||||  |  |  |
Db         51 GAPGRDGRD------GTPGEKGEKGDPGLIGPKGDIGETGVPGAEGPRGFPGIQGRKGEP 104

Qy         96 GECSVPPRSAFSAKRSESRVPPPSDAPLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAV 155
              || :   |||||    |: |  | : |: | ::  |:| |||  |||| | :||:||||
Db        105 GEGAYVYRSAFSV-GLETYVTIP-NMPIRFTKIFYNQQNHYDGSTGKFHCNIPGLYYFAY 162

Qy        156 HATVYRASLQFDLVKNGES-IASFFQFFGGWPKPASLSGGAMVRLEPEDQVWVQV-GVGD 213
              | |||   ::  | |  :: : :: |:       |  ||  :: ||  ||||:|| | |:
Db        163 HITVYMKDVKVSLFKKDKAMLFTYDQYQENNVDQA--SGSVLLHLEVGDQVWLQVYGEGE 220

Qy        214 YIGIYASIKTDSTFSGFLVYSD 235
                |:||    ||||:|||:| |
Db        221 RNGLYADNDNDSTFTGFLLYHD 242
```

```
RESULT 4
CA1A_HUMAN
ID   CA1A_HUMAN     STANDARD;      PRT;   680 AA.
AC   Q03692;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Collagen alpha 1(X) chain precursor.
GN   COL10A1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92109659; PubMed=1764025;
RA   Thomas J.T., Cresswell C.J., Rash B., Nicolai H., Jones T.,
RA   Solomon E., Grant M.E., Boot-Handford R.P.;
RT   "The human collagen X gene. Complete primary translated sequence and
RT   chromosomal localization.";
RL   Biochem. J. 280:617-623(1991).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93012005; PubMed=1397333;
RA   Reichenberger E., Beier F., Luvalle P., Olsen B.R., von der Mark K.,
RA   Bertling W.M.;
RT   "Genomic organization and full-length cDNA sequence of human collagen
RT   X.";
RL   FEBS Lett. 311:305-310(1992).
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Beier F., Lammi M.B., von der Mark K.;
RL   Submitted (JUN-1996) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE FROM N.A.
RA   Williams S.;
RL   Submitted (JAN-2001) to the EMBL/GenBank/DDBJ databases.
RN   [5]
RP   SEQUENCE OF 52-680 FROM N.A.
RX   MEDLINE=92267014; PubMed=1587271;
RA   Apte S.S., Seldin M.F., Hayashi M., Olsen B.R.;
RT   "Cloning of the human and mouse type X collagen genes and mapping of
RT   the mouse type X collagen gene to chromosome 10.";
RL   Eur. J. Biochem. 206:217-224(1992).
RN   [6]
RP   SEQUENCE OF 561-666 FROM N.A.
RX   MEDLINE=91243838; PubMed=2037056;
RA   Apte S., Mattei M.-G., Olsen B.R.;
RT   "Cloning of human alpha 1(X) collagen DNA and localization of the
RT   COL10A1 gene to the q21-q22 region of human chromosome 6.";
RL   FEBS Lett. 282:393-396(1991).
RN   [7]
RP   SEQUENCE OF 547-655 FROM N.A.
RX   MEDLINE=92077285; PubMed=1743401;
RA   Reichenberger E., Aigner T., von der Mark K., Stoeb H., Bertling W.;
RT   "In situ hybridization studies on the expression of type X collagen
RT   in fetal human cartilage.";
RL   Dev. Biol. 148:562-572(1991).
RN   [8]
RP   REVIEW ON VARIANTS.
RX   MEDLINE=97255959; PubMed=9101290;
RA   Kuivaniemi H., Tromp G., Prockop D.J.;
RT   "Mutations in fibrillar collagens (types I, II, III, and XI), fibril-
RT   associated collagen (type IX), and network-forming collagen (type X)
RT   cause a spectrum of diseases of bone, cartilage, and blood vessels.";
RL   Hum. Mutat. 9:300-315(1997).
RN   [9]
RP   VARIANTS SMCD ASP-598 AND PRO-614.
RX   MEDLINE=94136476; PubMed=8304336;
RA   Wallis G.A., Rash B., Sweetman W.A., Thomas J.T., Super M., Evans G.,
RA   Grant M.E., Boot-Handford R.P.;
```

```
RT   "Amino acid substitutions of conserved residues in the
RT   carboxyl-terminal domain of the alpha 1(X) chain of type X collagen
RT   occur in two unrelated families with metaphyseal chondrodysplasia
RT   type Schmid.";
RL   Am. J. Hum. Genet. 54:169-178(1994).
RN   [10]
RP   VARIANT SMCD ARG-591.
RX   MEDLINE=94272470; PubMed=8004099;
RA   McIntosh I., Abbott M.H., Warman M.L., Olsen B.R., Francomano C.A.;
RT   "Additional mutations of type X collagen confirm COL10A1 as the
RT   Schmid metaphyseal chondrodysplasia locus.";
RL   Hum. Mol. Genet. 3:303-307(1994).
RN   [11]
RP   VARIANT SMCD VAL-618.
RX   MEDLINE=95181449; PubMed=7876225;
RA   Chan D., Cole W.G., Rogers J.G., Bateman J.F.;
RT   "Type X collagen multimer assembly in vitro is prevented by a Gly618
RT   to Val mutation in the alpha 1(X) NC1 domain resulting in Schmid
RT   metaphyseal chondrodysplasia.";
RL   J. Biol. Chem. 270:4558-4562(1995).
RN   [12]
RP   VARIANTS SMCD ARG-545; GLU-595; HIS-597; LYS-617; ARG-644 AND GLY-648.
RX   MEDLINE=95331767; PubMed=7607655;
RA   Bonaventure J., Chaminade F., Maroteaux P.;
RT   "Mutations in three subdomains of the carboxy-terminal region of
RT   collagen type X account for most of the Schmid metaphyseal
RT   dysplasias.";
RL   Hum. Genet. 96:58-64(1995).
RN   [13]
RP   VARIANT SMCD PRO-600.
RX   MEDLINE=96375754; PubMed=8782043;
RA   Wallis G.A., Rash B., Sykes B., Bonaventure J., Maroteaux P.,
RA   Zabel B., Wynne-Davies R., Grant M.E., Boot-Handford R.P.;
RT   "Mutations within the gene encoding the alpha 1 (X) chain of type X
RT   collagen (COL10A1) cause metaphyseal chondrodysplasia type Schmid but
RT   not several other forms of metaphyseal chondrodysplasia.";
RL   J. Med. Genet. 33:450-457(1996).
RN   [14]
RP   VARIANTS SMCD GLU-18 AND ARG-18.
RX   MEDLINE=97220591; PubMed=9067753;
RA   Ikegawa S., Nakamura K., Nagano A., Haga N., Nakamura Y.;
RT   "Mutations in the N-terminal globular domain of the type X collagen
RT   gene (COL10A1) in patients with schmid metaphyseal
RT   chondrodysplasia.";
RL   Hum. Mutat. 9:131-135(1997).
RN   [15]
RP   VARIANTS SMD GLU-595.
RX   MEDLINE=99057503; PubMed=9837818;
RA   Ikegawa S., Nishimura G., Nagai T., Hasegawa T., Ohashi H.,
RA   Nakamura Y.;
RT   "Mutation of the type X collagen gene 'COL10A1' causes
RT   spondylometaphyseal dysplasia.";
RL   Am. J. Hum. Genet. 63:1659-1662(1998).
RN   [16]
RP   VARIANT SMCD CYS-597.
RX   MEDLINE=99069781; PubMed=9852679;
RA   Sawai H., Ida A., Nakata Y., Koyama K.;
RT   "Novel missense mutation resulting in the substitution of tyrosine by
RT   cysteine at codon 597 of the type X collagen gene associated with
RT   Schmid metaphyseal chondrodysplasia.";
RL   J. Hum. Genet. 43:259-261(1998).
CC   -!- FUNCTION: TYPE X COLLAGEN IS A PRODUCT OF HYPERTHROPHIC
CC       CHONDROTOCYTES AND HAS BEEN LOCALIZED TO PRESUMPTIVE
CC       MINERALIZATION ZONES OF HYALINE CARTILAGE.
CC   -!- SUBUNIT: HOMOTRIMER.
CC   -!- PTM: PROLINES AT THE THIRD POSITION OF THE TRIPEPTIDE REPEATING
CC       UNIT (G-X-Y) ARE HYDROXYLATED IN SOME OR ALL OF THE CHAINS.
CC   -!- DISEASE: DEFECTS IN COL10A1 ARE THE CAUSE OF SCHMID TYPE
CC       METAPHYSEAL CHONDRODYSPLASIA (SMCD), A DOMINANTLY INHERITED
CC       DISORDER OF THE OSSEOUS SKELETON. THE CARDINAL FEATURES OF THE
CC       PHENOTYPE ARE MILD SHORT STATURE, COXA VARA AND A WADDLING GAIT.
CC       RADIOGRAPHY USUALLY SHOWS SCLEROSIS OF THE RIBS, FLARING OF THE
CC       METAPHYSES, AND A WIDE IRREGULAR GROWTH PLATE, ESPECIALLY OF THE
CC       KNEES.
CC   -!- DISEASE: DEFECTS IN COL10A1 ARE THE CAUSE OF SPONDYLOMETAPHYSEAL
CC       DYSPLASIA (SMD). SMD COMPRISES A HETEROGENEOUS GROUP OF HERITABLE
CC       SKELETAL DYSPLASIAS CHARACTERIZED BY MODIFICATIONS OF THE
CC       VERTEBRAL BODIES OF THE SPINE AND METAPHYSES OF THE TUBULAR BONES.
CC   -!- SIMILARITY: STRONG, TO ALPHA 1 AND 2 TYPE VIII COLLAGENS.
CC   -!- SIMILARITY: CONTAINS 1 C1Q DOMAIN.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X60382; CAA42933.1; -.
DR   EMBL; X65120; CAA46236.1; -.
DR   EMBL; X98568; CAA67178.1; -.
DR   EMBL; AL121963; CAB87590.1; -.
DR   EMBL; S68531; AAC60615.1; -.
DR   EMBL; X58879; CAA41686.1; -.
DR   EMBL; M74050; AAA61221.1; -.
DR   EMBL; X72579; CAA51170.1; -.
DR   EMBL; X72580; CAA51170.1; JOINED.
DR   PIR; S15826; S15826.
DR   PIR; S30086; S30086.
DR   PIR; A43901; A43901.
DR   PIR; S18249; S18249.
DR   PIR; S21856; S21856.
DR   PIR; S26396; S26396.
DR   Genew; HGNC:2185; COL10A1.
DR   MIM; 120110; -.
DR   MIM; 156500; -.
DR   MIM; 184250; -.
DR   InterPro; IPR001073; C1q.
DR   InterPro; IPR000087; Collagen.
DR   Pfam; PF00386; C1q; 1.
DR   Pfam; PF01391; Collagen; 8.
DR   PRINTS; PR00007; COMPLEMNTC1Q.
DR   SMART; SM00110; C1Q; 1.
DR   PROSITE; PS01113; C1Q; 1.
KW   Extracellular matrix; Connective tissue; Repeat; Hydroxylation;
KW   Cartilage; Collagen; Signal; Disease mutation; Polymorphism.
FT   SIGNAL        1     18       POTENTIAL.
FT   CHAIN        19    680       COLLAGEN ALPHA 1(X) CHAIN.
FT   DOMAIN       19     56       NONHELICAL REGION (NC2).
FT   DOMAIN       57    519       TRIPLE-HELICAL REGION.
FT   DOMAIN      520    680       NONHELICAL REGION (NC1).
FT   DOMAIN      545    680       C1Q.
FT   VARIANT      18     18       G -> E (IN SMCD).
FT                                /FTId=VAR_001838.
FT   VARIANT      18     18       G -> R (IN SMCD).
FT                                /FTId=VAR_001839.
FT   VARIANT     545    545       G -> R.
FT                                /FTId=VAR_001840.
FT   VARIANT     591    591       C -> R (IN SMCD).
FT                                /FTId=VAR_001841.

  Query Match           31.6%;  Score 418.5;  DB 1;  Length 680;
  Best Local Similarity 34.9%;  Pred. No. 3.5e-22;
  Matches   98;  Conservative   34;  Mismatches   82;  Indels   67;  Gaps    6;

Qy        15 GSPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPGAPGEKGEGGRPGLP 74
             |:| |   |      | ||||| |  |::|:|| :|  |  |||| || :|  | ||:|
Db       401 GNPGLPGPKGDPGVGGPPGLPGPVGPAGAKGMPGHNGEAGPRGAPGIPGTRGPIGPPGIP 460

Qy        75 ----------------------------GPRGDPGPRGEA------GPAGPTGPAGECS 99
                                          || | ||||| :      || || || |:
Db       461 GFPGSKGDPGSPGPPGPAGIATKGLNGPTGPPGPPGPRGHSGEPGLPGPPGPPGPPGQAV 520

Qy       100 VP----------------------------PRSAFSAKRSESRVPPPSDAPLPFDRVLVNE 132
               :|                           | |||:   |::   |    |:|||::| |
```

```
Db        521 MPEGFIKAGQRPSLSGTPLVSANQGVTGMPVSAFTVILSKAY--PAIGTPIPFDKILYNR 578

Qy        133 QGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGESIA-SFFQFFGGWPKPASL 191
              | |||  || ||||:||:|||: |  |     :   | |||  :  :: ::  |:   |
Db        579 QQHYDPRTGIFTCQIPGIYYFSYHVHVKGTHVWVGLYKNGTPVMYTYDEYTKGYLDQA-- 636

Qy        192 SGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLV 232
              || |:: |   ||||:|:   :  |:|:|    |:||||||
Db        637 SGSAIIDLTENDQVWLQLPNAESNGLYSSEYVHSSFSGFLV 677


RESULT 5
CA1A_CHICK
ID   CA1A_CHICK     STANDARD;      PRT;   674 AA.
AC   P08125;
DT   01-AUG-1988 (Rel. 08, Created)
DT   01-NOV-1991 (Rel. 20, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   Collagen alpha 1(X) chain precursor.
GN   COL10A1.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE OF 48-674 FROM N.A., AND SEQUENCE OF 103-117 AND 453-466.
RX   MEDLINE=86168227; PubMed=3082876;
RA   Ninomiya Y., Gordon M., van der Rest M., Schmid T., Linsenmayer T.,
RA   Olsen B.R.;
RT   "The developmentally regulated type X collagen gene contains a long
RT   open reading frame without introns.";
RL   J. Biol. Chem. 261:5041-5050(1986).
RN   [2]
RP   SEQUENCE OF 1-75 FROM N.A.
RX   MEDLINE=89054019; PubMed=2461368;
RA   Luvalle P., Ninomiya Y., Rosenblum N.D., Olsen B.R.;
RT   "The type X collagen gene. Intron sequences split the 5'-untranslated
RT   region and separate the coding regions for the non-collagenous amino-
RT   terminal and triple-helical domains.";
RL   J. Biol. Chem. 263:18378-18385(1988).
RN   [3]
RP   REVISIONS TO C-TERMINUS.
RX   MEDLINE=89380199; PubMed=2476437;
RA   Yamaguchi N., Benya P.D., van der Rest M., Ninomiya Y.;
RT   "The cloning and sequencing of alpha 1(VIII) collagen cDNAs
RT   demonstrate that type VIII collagen is a short chain collagen and
RT   contains triple-helical and carboxyl-terminal non-triple-helical
RT   domains similar to those of type X collagen.";
RL   J. Biol. Chem. 264:16022-16029(1989).
CC   -!- FUNCTION: TYPE X COLLAGEN IS A PRODUCT OF HYPERTHROPHIC
CC       CHONDROTOCYTES AND HAS BEEN LOCALIZED TO PRESUMPTIVE
CC       MINERALIZATION ZONES OF HYALINE CARTILAGE.
CC   -!- SUBUNIT: HOMOTRIMER.
CC   -!- PTM: PROLINES AT THE THIRD POSITION OF THE TRIPEPTIDE REPEATING
CC       UNIT (G-X-Y) ARE HYDROXYLATED IN SOME OR ALL OF THE CHAINS.
CC   -!- SIMILARITY: STRONG, TO ALPHA 1 AND 2 TYPE VIII COLLAGENS.
CC   -!- SIMILARITY: CONTAINS 1 C1Q DOMAIN.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M13496; AAA48736.1; ALT_SEQ.
DR   EMBL; J04194; AAA48634.1; -.
DR   PIR; A31896; A31896.
DR   InterPro; IPR001073; C1q.
DR   InterPro; IPR000087; Collagen.
DR   Pfam; PF00386; C1q; 1.
DR   Pfam; PF01391; Collagen; 8.
DR   PRINTS; PR00007; COMPLEMNTC1Q.
DR   SMART; SM00110; C1Q; 1.
DR   PROSITE; PS01113; C1Q; 1.
KW   Extracellular matrix; Connective tissue; Repeat; Hydroxylation;
KW   Cartilage; Collagen; Signal.
FT   SIGNAL        1     18
FT   CHAIN        19    674       COLLAGEN ALPHA 1(X) CHAIN.
FT   DOMAIN       19     52       NONHELICAL REGION (NC2).
FT   DOMAIN       53    512       TRIPLE-HELICAL REGION.
FT   DOMAIN      513    674       NONHELICAL REGION (NC1).
FT   DOMAIN      539    674       C1Q.
FT   MOD_RES     453    453       HYDROXYLATION.
FT   MOD_RES     456    456       HYDROXYLATION.
SQ   SEQUENCE   674 AA;  66434 MW;  EAB48B1EF174B145 CRC64;

  Query Match           31.5%;  Score 417;  DB 1;  Length 674;
  Best Local Similarity  36.0%;  Pred. No. 4.4e-22;
  Matches   96;  Conservative   34;  Mismatches   84;  Indels   53;  Gaps    6;

Qy         14 AGSPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPGAPGEKGEGGRPGL 73
                || | |      |    | ||: | ||  |  |:||  |  |  | ||||| ||| | |||
Db        408 AGHPGLPGPVGPQGVKGVPGINGEPGPRGPSGIPGVRGPIGPPGMPGAPGAKGEAGAPGL 467

Qy         74 PGPR--------------GDPGPRGEA------GPAGPTGPAGECSVP------------ 101
              |||               | |||:| :      || || || |: ::|
Db        468 PGPAGIVTKGLRGPMGPLGPPGPKGNSGEPGLPGPPGPPGPPGQSTIPEGYVKGESRELS 527

Qy        102 ---------------PRSAFSAKRSESRVPPPSDAPLPFDRVLVNEQGHYDAVTGKFTC 145
                            | |||:   |::   | :  |: ||::| | | |||  || |||
Db        528 GMSFMKAGANQALTGMPVSAFTVILSKAY--PGATVPIKFDKILYNRQQHYDPRTGIFTC 585

Qy        146 QVPGVYYFAVHATVYRASLQFDLVKNGESIA-SFFQFFGGWPKPASLSGGAMVRLEPEDQ 204
              ::||:|||: |      ::   | |||  :  :: ::  |:   |  || |:: |   ||
Db        586 RIPGLYYFSYHVHAKGTNVWVALYKNGSPVMYTYDEYQKGYLDQA--SGSAVIDLMENDQ 643

Qy        205 VWVQVGVGDYIGIYASIKTDSTFSGFL 231
              ||:|:   :  |:|:|    |:|||||
Db        644 VWLQLPNSESNGLYSSEYVHSSFSGFL 670


RESULT 6
C1QC_MOUSE
ID   C1QC_MOUSE     STANDARD;      PRT;   246 AA.
AC   Q02105;
DT   01-JUL-1993 (Rel. 26, Created)
DT   01-JUL-1993 (Rel. 26, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Complement C1q subcomponent, C chain precursor.
GN   C1QG OR C1QC.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ICR; TISSUE=Macrophage;
RX   MEDLINE=93011118; PubMed=1396691;
RA   Petry F., Reid K.B.M., Loos M.;
RT   "Isolation, sequence analysis and characterization of cDNA clones
RT   coding for the C chain of mouse C1q. Sequence similarity of
RT   complement subcomponent C1q, collagen type VIII and type X and
RT   precerebellin.";
RL   Eur. J. Biochem. 209:129-134(1992).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BALB/c; TISSUE=Liver;
RX   MEDLINE=96186528; PubMed=8606057;
RA   Petry F., McClive P.J., Botto M., Morley B.J., Morahan G., Loos M.;
RT   "The mouse C1q genes are clustered on chromosome 4 and show
RT   conservation of gene organization.";
RL   Immunogenetics 43:370-376(1996).
CC   -!- FUNCTION: C1Q ASSOCIATES WITH THE PROENZYMES C1R AND C1S TO YIELD
```

```
CC       C1, THE FIRST COMPONENT OF THE SERUM COMPLEMENT SYSTEM. THE
CC       COLLAGEN-LIKE REGIONS OF C1Q INTERACT WITH THE CA(2+)-DEPENDENT
CC       C1R(2)C1S(2) PROENZYME COMPLEX, AND EFFICIENT ACTIVATION OF C1
CC       TAKES PLACE ON INTERACTION OF THE GLOBULAR HEADS OF C1Q WITH THE
CC       FC REGIONS OF IGG OR IGM ANTIBODY PRESENT IN IMMUNE COMPLEXES.
CC   -!- SUBUNIT: C1 IS A CALCIUM-DEPENDENT TRIMOLECULAR COMPLEX OF C1Q, R
CC       AND S IN THE MOLAR RATION OF 1:2:2. C1Q SUBCOMPONENT IS COMPOSED
CC       OF NINE SUBUNITS, SIX OF WHICH ARE DISULFIDE-LINKED DIMERS OF THE
CC       A AND B CHAINS, AND THREE OF WHICH ARE DISULFIDE-LINKED DIMERS OF
CC       THE C CHAIN.
CC   -!- SIMILARITY: CONTAINS 1 COLLAGENOUS DOMAIN.
CC   -!- SIMILARITY: CONTAINS 1 C1Q DOMAIN.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X66295; CAA46993.1; -.
DR   EMBL; X92960; CAA63535.1; -.
DR   PIR; S29328; S29328.
DR   MGD; MGI:88225; C1qc.
DR   InterPro; IPR001073; C1q.
DR   InterPro; IPR000087; Collagen.
DR   Pfam; PF00386; C1q; 1.
DR   Pfam; PF01391; Collagen; 1.
DR   PRINTS; PR00007; COMPLEMNTC1Q.
DR   SMART; SM00110; C1Q; 1.
DR   PROSITE; PS01113; C1Q; 1.
KW   Complement pathway; Plasma; Hydroxylation; Glycoprotein; Collagen;
KW   Repeat; Signal.
FT   SIGNAL        1     29
FT   CHAIN        30    246       COMPLEMENT C1Q SUBCOMPONENT, C CHAIN.
FT   DOMAIN       32    113       COLLAGEN-LIKE.
FT   DOMAIN      114    246       C1Q.
FT   DISULFID     33     33       INTERCHAIN (WITH OTHER C CHAIN)
FT                                (BY SIMILARITY).
FT   MOD_RES      37     37       HYDROXYLATION (BY SIMILARITY).
FT   MOD_RES      40     40       HYDROXYLATION (BY SIMILARITY).
FT   MOD_RES      43     43       HYDROXYLATION (BY SIMILARITY).
FT   MOD_RES      61     61       HYDROXYLATION (BY SIMILARITY).
FT   MOD_RES      64     64       HYDROXYLATION (BY SIMILARITY).
FT   MOD_RES      73     73       HYDROXYLATION (BY SIMILARITY).
FT   MOD_RES      76     76       HYDROXYLATION (BY SIMILARITY).
FT   MOD_RES      79     79       HYDROXYLATION (BY SIMILARITY).
FT   MOD_RES      82     82       HYDROXYLATION (BY SIMILARITY).
FT   MOD_RES      97     97       HYDROXYLATION (BY SIMILARITY).
FT   MOD_RES     100    100       HYDROXYLATION (BY SIMILARITY).
FT   MOD_RES     106    106       HYDROXYLATION (BY SIMILARITY).
FT   MOD_RES     109    109       HYDROXYLATION (BY SIMILARITY).
SQ   SEQUENCE   246 AA;  25966 MW;  2F79EA1274BCB8E0 CRC64;

  Query Match          31.1%;  Score 412;  DB 1;  Length 246;
  Best Local Similarity  42.2%;  Pred. No. 3.7e-22;
  Matches  103;  Conservative   33;  Mismatches   84;  Indels   24;  Gaps   10;

Qy          4 LLVLLLLGLAAGSPPLDDNKIPSLC---PGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPG 60
              ||:| || |     ||  ::  : |   || ||:|| ||  |  || |  |  |    ||
Db         15 LLLLFLLAL-----PL-RSQASAGCYGIPGMPGMPGAPGKDGHDGLQGPKGEPGIPAVPG 68

Qy         61 APGEKGEGGRPGLPGPRGDPGPRGEA---GPAGPTGPAGECSVPPR------SAFSAKRS 111
                | ||: | ||:|| ||  |||| :   |  || || ||  |  |       | |:  |
Db         69 TQGPKGQKGEPGMPGHRGKNGPRGTSGLPGDPGPRGPPGEPGVEGRYKQKHQSVFTVTRQ 128

Qy        112 ESRVPPPSDAPLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKN 171
               ::  | ::| : |: |: | ||||:  ||||||:|||:||| |: | : |:|   | |
Db        129 TTQY-PEANALVRFNSVVTNPQGHYNPSTGKFTCEVPGLYYF-VYYTSHTANLCVHLNLN 186

Qy        172 GESIASFFQFFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFL 231
              :|||        |  | ||||::||:  |:||  : | || |:     ::| |||||

Db        187 LARVASFCDHMFN-SKQVS-SGGALLRLQRGDEVW--LSVNDYNGMVGIEGSNSVFSGFL 242

Qy        232 VYSD 235
              :: |
Db        243 LFPD 246


RESULT 7
CA1A_BOVIN
ID   CA1A_BOVIN     STANDARD;      PRT;   674 AA.
AC   P23206;
DT   01-NOV-1991 (Rel. 20, Created)
DT   01-NOV-1991 (Rel. 20, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   Collagen alpha 1(X) chain precursor.
GN   COL10A1.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Cartilage;
RX   MEDLINE=91113131; PubMed=1703407;
RA   Thomas J.T., Kwan A.P.L., Grant M.E., Boot-Handford R.P.;
RT   "Isolation of cDNAs encoding the complete sequence of bovine type X
RT   collagen. Evidence for the condensed nature of mammalian type X
RT   collagen genes.";
RL   Biochem. J. 273:141-148(1991).
CC   -!- FUNCTION: TYPE X COLLAGEN IS A PRODUCT OF HYPERTHROPHIC
CC       CHONDROTOCYTES AND HAS BEEN LOCALIZED TO PRESUMPTIVE
CC       MINERALIZATION ZONES OF HYALINE CARTILAGE.
CC   -!- SUBUNIT: HOMOTRIMER.
CC   -!- PTM: PROLINES AT THE THIRD POSITION OF THE TRIPEPTIDE REPEATING
CC       UNIT (G-X-Y) ARE HYDROXYLATED IN SOME OR ALL OF THE CHAINS.
CC   -!- SIMILARITY: STRONG, TO ALPHA 1 AND 2 TYPE VIII COLLAGENS.
CC   -!- SIMILARITY: CONTAINS 1 C1Q DOMAIN.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X53556; CAA37624.1; -.
DR   PIR; S13301; S13301.
DR   InterPro; IPR001073; C1q.
DR   InterPro; IPR000087; Collagen.
DR   Pfam; PF00386; C1q; 1.
DR   Pfam; PF01391; Collagen; 9.
DR   PRINTS; PR00007; COMPLEMNTC1Q.
DR   ProDom; PD000007; Collagen; 1.
DR   SMART; SM00110; C1Q; 1.
DR   PROSITE; PS01113; C1Q; 1.
KW   Extracellular matrix; Connective tissue; Repeat; Hydroxylation;
KW   Cartilage; Collagen; Signal; Glycoprotein.
FT   SIGNAL        1     18
FT   CHAIN        19    674       COLLAGEN ALPHA 1(X) CHAIN.
FT   DOMAIN       19     56       NONHELICAL REGION (NC2).
FT   DOMAIN       57    519       TRIPLE-HELICAL REGION.
FT   DOMAIN      520    674       NONHELICAL REGION (NC1).
FT   DOMAIN      539    674       C1Q.
FT   DISULFID    194    197       BY SIMILARITY.
FT   MOD_RES     460    460       HYDROXYLATION (BY SIMILARITY).
FT   MOD_RES     463    463       HYDROXYLATION (BY SIMILARITY).
FT   CARBOHYD    611    611       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   674 AA;  65546 MW;  CD4CA73A03E004CA CRC64;

  Query Match          31.1%;  Score 411.5;  DB 1;  Length 674;
  Best Local Similarity  34.9%;  Pred. No. 1.1e-21;
```

```
  Matches   96;  Conservative   34;  Mismatches   84;  Indels   61;  Gaps    6;

Qy         15 GSPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPGAPGEKGEGGRPGLP 74
              |:| |   |      | |||||  |  |::|:|| :|  |  | || || :|  | ||:|
Db        401 GNPGLPGPKGDPGIAGSPGLPGPVGPAGAKGVPGHNGEAGPRGVPGIPGTRGPIGPPGIP 460

Qy         75 ----------------------------GPRGDPGPRGEA------GPAGPTGPAGECS 99
                                          || | ||||| |      || || || |: : .
Db        461 GFPGSKGDVGTPGPPGPAGIAVKGLNGPTGPPGPPGPRGNAGEPGLPGPPGPPGPPGQVA 520

Qy         100 VP--------------------PRSAFSAKRSESRVPPPSDAPLPFDRVLVNEQGHYDA 138
               :|                    | |||:   |::   |    |:|||::| |:| |||
Db         521 LPEDFVKAGQRPFVSANQGVTGMPVSAFTVILSKAY--PAIGTPIPFDKILYNKQQHYDP 578

Qy         139 VTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGESIA-SFFQFFGGWPKPASLSGGAMV 197
               || |||::||:|||: |  |        | |||  :  :: ::  |:   |  || |::
Db         579 RTGIFTCKIPGIYYFSYHIHVKGTHAWVGLYKNGTPVMYTYDEYIKGYLDQA--SGSAVI 636

Qy         198 RLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLV 232
                 |  ||||:|:      |:|:     |:||||||
Db         637 DLTENDQVWLQLPNAGSNGLYSPEYVHSSFSGFLV 671


RESULT 8
CQT7_HUMAN
ID   CQT7_HUMAN     STANDARD;      PRT;   289 AA.
AC   Q9BXJ2;
DT   15-JUN-2002 (Rel. 41, Created)
DT   15-JUN-2002 (Rel. 41, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Complement-c1q tumor necrosis factor-related protein 7 precursor.
GN   C1QTNF7 OR CTRP7.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Piddington C.S., Sheppard P.O., Bishop P., Lasser G.W.;
RT   "Homo sapiens complement-c1q tumor necrosis factor-related protein.";
RL   Submitted (DEC-2000) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Testis;
RA   Strausberg R.;
RL   Submitted (JAN-2002) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: CONTAINS 1 COLLAGENOUS DOMAIN.
CC   -!- SIMILARITY: CONTAINS 1 C1Q DOMAIN.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF329839; AAK17963.1; -.
DR   EMBL; BC022187; AAH22187.1; -.
DR   Genew; HGNC:14342; C1QTNF7.
DR   InterPro; IPR001073; C1q.
DR   InterPro; IPR000087; Collagen.
DR   Pfam; PF00386; C1q; 1.
DR   Pfam; PF01391; Collagen; 2.
DR   PRINTS; PR00007; COMPLEMNTC1Q.
DR   SMART; SM00110; C1Q; 1.
DR   PROSITE; PS01113; C1Q; 1.
KW   Collagen; Signal.
FT   SIGNAL       1     16       POTENTIAL.
FT   CHAIN       17    289       COMPLEMENT-C1Q TUMOR NECROSIS FACTOR-
FT                               RELATED PROTEIN 7.
FT   DOMAIN      38    139       COLLAGEN-LIKE.
FT   DOMAIN     141    276       C1Q.

SQ   SEQUENCE   289 AA;  30683 MW;  A61609FF86D26946 CRC64;

  Query Match           31.0%;  Score 410.5;  DB 1;  Length 289;
  Best Local Similarity   40.0%;  Pred. No. 5.4e-22;
  Matches   98;  Conservative   26;  Mismatches   84;  Indels   37;  Gaps    7;

Qy         24 IPSL--CPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAP--------------GAPGEKG 66
              || |   || ||  |:|| ||  ||||||||||| |                |  ||||
Db         36 IPGLPGPPGPPGANGSPGPHGRIGLPGRDGRDGRKGEKGEKGTAGLRGKTGPLGLAGEKG 95

Qy         67 EGGRPGLPGPRGDPGPRGEAGPAGPTGPAGE-------------C-SVPPRSAFSAKRS 111
              : |  |  || |  | :|| || || || ||:                | |:  :|||   :
Db         96 DQGETGKKGPIGPEGEKGEVGPIGPPGPKGDRGEQGDPGLPGVCRCGSIVLKSAFSVGIT 155

Qy        112 ESRVPPPSDAPLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKN 171
                |  |   |: |::|| ||  ||:  |||| |  ||:|||:   |:    |  || |
Db        156 TSY--PEERLPIIFNKVLFNEGEHYNPATGKFICAFPGIYYFSYDITLANKHLAIGLVHN 213

Qy        172 GESIASFFQFFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIK-TDSTFSGF 230
              |:     |    |   | ||  :: |:|||:||:::   |  |:::    || ||||
Db        214 GQYRIKTFDANTGNHDVA--SGSTVIYLQPEDEVWLEIFFTDQNGLFSDPGWADSLFSGF 271

Qy        231 LVYSD 235
              |:| |
Db        272 LLYVD 276


RESULT 9
CA1A_MOUSE
ID   CA1A_MOUSE     STANDARD;      PRT;   680 AA.
AC   Q05306;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   01-NOV-1995 (Rel. 32, Last annotation update)
DE   Collagen alpha 1(X) chain precursor.
GN   COL10A1.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BALB/c;
RX   MEDLINE=93143676; PubMed=8424763;
RA   Elima K., Eerola I., Rosati R., Metsaranta M., Garofalo S., Perala M.,
RA   de Crombrugghe B., Vuorio E.;
RT   "The mouse collagen X gene: complete nucleotide sequence, exon
RT   structure and expression pattern.";
RL   Biochem. J. 289:247-253(1993).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=129/Sv; TISSUE=Liver;
RX   MEDLINE=93238750; PubMed=8477738;
RA   Kong R.Y.C., Kwan K.M., Lau E.T., Thomas J.T., Boot-Handford R.P.,
RA   Grant M.E., Cheah K.S.E.;
RT   "Intron-exon structure, alternative use of promoter and expression of
RT   the mouse collagen X gene, Col10a-1.";
RL   Eur. J. Biochem. 213:99-111(1993).
RN   [3]
RP   SEQUENCE OF 51-680 FROM N.A.
RC   STRAIN=DBA/2J;
RX   MEDLINE=92267014; PubMed=1587271;
RA   Apte S.S., Seldin M.F., Hayashi M., Olsen B.R.;
RT   "Cloning of the human and mouse type X collagen genes and mapping of
RT   the mouse type X collagen gene to chromosome 10.";
RL   Eur. J. Biochem. 206:217-224(1992).
RN   [4]
RP   SEQUENCE OF 385-627 FROM N.A.
RC   STRAIN=C57BL/6;
RX   MEDLINE=92182017; PubMed=1543751;
RA   Elima K., Metsaeranta M., Kallio J., Peraelae M., Eerola I.,
RA   Garofalo S., de Crombrugghe B., Vuorio E.;
RT   "Specific hybridization probes for mouse alpha 2(IX) and alpha 1(X)
```

```
RT   collagen mRNAs.";
RL   Biochim. Biophys. Acta 1130:78-80(1992).
CC   -!- FUNCTION: TYPE X COLLAGEN IS A PRODUCT OF HYPERTHROPHIC
CC       CHONDROTOCYTES AND HAS BEEN LOCALIZED TO PRESUMPTIVE
CC       MINERALIZATION ZONES OF HYALINE CARTILAGE.
CC   -!- SUBUNIT: HOMOTRIMER.
CC   -!- PTM: PROLINES AT THE THIRD POSITION OF THE TRIPEPTIDE REPEATING
CC       UNIT (G-X-Y) ARE HYDROXYLATED IN SOME OR ALL OF THE CHAINS.
CC   -!- SIMILARITY: STRONG, TO ALPHA 1 AND 2 TYPE VIII COLLAGENS.
CC   -!- SIMILARITY: CONTAINS 1 C1Q DOMAIN.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X67348; CAA47763.1; -.
DR   EMBL; X65121; CAA46237.1; -.
DR   EMBL; X63013; CAA44741.1; -.
DR   EMBL; Z21610; CAA79736.1; -.
DR   PIR; S28807; S28807.
DR   PIR; S31216; S31216.
DR   PIR; S22215; S22215.
DR   MGD; MGI:88445; Col10a1.
DR   InterPro; IPR001073; C1q.
DR   InterPro; IPR000087; Collagen.
DR   Pfam; PF00386; C1q; 1.
DR   Pfam; PF01391; Collagen; 9.
DR   PRINTS; PR00007; COMPLEMNTC1Q.
DR   ProDom; PD000007; Collagen; 2.
DR   SMART; SM00110; C1Q; 1.
DR   PROSITE; PS01113; C1Q; 1.
KW   Extracellular matrix; Connective tissue; Repeat; Hydroxylation;
KW   Cartilage; Collagen; Signal.
FT   SIGNAL        1     18       POTENTIAL.
FT   CHAIN        19    680       COLLAGEN ALPHA 1(X) CHAIN.
FT   DOMAIN       19     56       NONHELICAL REGION (NC2).
FT   DOMAIN       57    519       TRIPLE-HELICAL REGION.
FT   DOMAIN      520    680       NONHELICAL REGION (NC1).
FT   DOMAIN      545    680       C1Q.
FT   CONFLICT    248    248       P -> L (IN REF. 3).
FT   CONFLICT    286    286       A -> S (IN REF. 2).
FT   CONFLICT    306    306       L -> F (IN REF. 3).
FT   CONFLICT    417    417       T -> S (IN REF. 3).
FT   CONFLICT    451    451       R -> K (IN REF. 4).
FT   CONFLICT    500    500       H -> L (IN REF. 3).
FT   CONFLICT    567    567       A -> C (IN REF. 3).
FT   CONFLICT    569    569       I -> H (IN REF. 3).
FT   CONFLICT    571    572       FD -> IY (IN REF. 3).
FT   CONFLICT    635    635       Q -> T (IN REF. 3).
SQ   SEQUENCE   680 AA;  66775 MW;  FE984CA99AF708E2 CRC64;

  Query Match           30.8%;  Score 408;  DB 1;  Length 680;
  Best Local Similarity  33.4%;  Pred. No. 1.9e-21;
  Matches   99;  Conservative   34;  Mismatches   75;  Indels   88;  Gaps    9;

Qy        15 GSPPLDDNKIPSLCPGHPGLP----------------------GTPGHH------GSQ 44
             | | |:  |      |:||||                      | |||:      |
Db       392 GEPGLNGPK------GNPGLPGQKGDPGVGGTPGLRGPVGPVGAKGVPGHNGEAGPRGEP 445

Qy        45 GLPGRDGRDGRDGAPGAPGEKGEGGRPGLPGPR-------------GDPGPRGEA---- 86
             |:||  |  |  | || || ||: | || |||              | |||||  :
Db       446 GIPGTRGPTGPPGVPGFPGSKGDPGNPGAPGPAGIATKGLNGPTGPPGPPGPRGHSGEPG 505

Qy        87 --GPAGPTGPAGECSVP-------------------------PRSAFSAKRSESRVPP 117
               || || || |:  :|                        | |||:   |::   |
Db       506 LPGPPGPPGPPGQAVMPDGFIKAGQRPRLSGMPLVSANHGVTGMPVSAFTVILSKAY--P 563

Qy       118 PSDAPLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNG-ESIA 176
               ||:||| :| | | |||  :| |||::||:|||: |  |    :   | |||  ::
Db       564 AVGAPIPFDEILYNRQQHYDPRSGIFTCKIPGIYYFSYHVHVKGTHVWVGLYKNGTPTMY 623

Qy       177 SFFQFFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLV 232
             :: ::  |:   |  || |:: |   ||||:|:    :  |:|:|    |:||||||
Db       624 TYDEYSKGYLDQA--SGSAIMELTENDQVWLQLPNAESNGLYSSEYVHSSFSGFLV 677


RESULT 10
CQT2_HUMAN
ID   CQT2_HUMAN     STANDARD;      PRT;   285 AA.
AC   Q9BXJ5;
DT   15-JUN-2002 (Rel. 41, Created)
DT   15-JUN-2002 (Rel. 41, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Complement-c1q tumor necrosis factor-related protein 2 precursor.
GN   C1QTNF2 OR CTRP2.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Piddington C.S., Bishop P.;
RT   "Homo sapiens complement-c1q tumor necrosis factor-related protein.";
RL   Submitted (DEC-2000) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Muscle;
RA   Strausberg R.;
RL   Submitted (JUL-2001) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: CONTAINS 1 COLLAGENOUS DOMAIN.
CC   -!- SIMILARITY: CONTAINS 1 C1Q DOMAIN.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF329836; AAK17960.1; -.
DR   EMBL; BC011699; AAH11699.1; -.
DR   Genew; HGNC:14325; C1QTNF2.
DR   InterPro; IPR001073; C1q.
DR   InterPro; IPR000087; Collagen.
DR   Pfam; PF00386; C1q; 1.
DR   Pfam; PF01391; Collagen; 2.
DR   PRINTS; PR00007; COMPLEMNTC1Q.
DR   SMART; SM00110; C1Q; 1.
DR   PROSITE; PS01113; C1Q; 1.
KW   Collagen; Signal.
FT   SIGNAL        1     15       POTENTIAL.
FT   CHAIN        16    285       COMPLEMENT-C1Q TUMOR NECROSIS FACTOR-
FT                                RELATED PROTEIN 2.
FT   DOMAIN       40    141       COLLAGEN-LIKE.
FT   DOMAIN      143    285       C1Q.
SQ   SEQUENCE   285 AA;  29952 MW;  7E31FF9868D4EDFA CRC64;

  Query Match           30.4%;  Score 403;  DB 1;  Length 285;
  Best Local Similarity  36.3%;  Pred. No. 1.8e-21;
  Matches   98;  Conservative   29;  Mismatches   91;  Indels   52;  Gaps    8;

Qy         9 LLGLAA------GSPPLDDNKIPSLC--------PGHPGLPGTPGHHGSQGLPGRDGRDG 54
             |||  |      ||| |       :|        || || ||  |  |  | ||:||:||
Db        18 LLGAFARRDFRKGSPQL-------VCSLPGPQGPPGPPGAPGPSGMMGRMGFPGKDGQDG 70

Qy        55 RDGAPGAPGEKGEGGRP---GLPGPRGDPGPRGEAGPAGP---TGPAGECSVPPRSAFSA 108
               ||  |  ||:|  ||    | |||:|  |  | ||| ||    |  |:   | :
Db        71 HDGDRGDSGEEGPPGRTGNRGKPGPKGKAGAIGRAGPRGPKGVNGTPGKHGTPGKKGPKG 130

Qy       109 KRSESRVPPPSDA---------------------PLPFDRVLVNEQGHYDAVTGKFTCQ 146
             |: |  :| |                        |: ||::|:|| |||:| :||| |
```

```
Db        131 KKGEPGLPGPCSCGSGHTKSAFSVAVTKSYPRERLPIKFDKILMNEGGHYNASSGKFVCG 190

Qy        147 VPGVYYFAVHATVYRASLQFDLVKNGESIASFFQFFGGWPKPASLSGGAMVRLEPEDQVW 206
              |||:|||    |:    |   || ||:     |    |    |  ||  :: |:  |:||
Db        191 VPGIYYFTYDITLANKHLAIGLVHNGQYRIRTFDANTGNHDVA--SGSTILALKQGDEVW 248

Qy        207 VQVGVGDYIGI-YASIKTDSTFSGFLVYSD 235
              :|:   :  |: |    ||| |:|||:|:|
Db        249 LQIFYSEQNGLFYDPYWTDSLFTGFLIYAD 278


RESULT 11
C1QC_HUMAN
ID   C1QC_HUMAN     STANDARD;      PRT;   245 AA.
AC   P02747; Q96DL2; Q96H05;
DT   21-JUL-1986 (Rel. 01, Created)
DT   15-JUN-2002 (Rel. 41, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Complement C1q subcomponent, C chain precursor.
GN   C1QG OR C1QC.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Monocytes;
RX   MEDLINE=91174759; PubMed=1706597;
RA   Sellar G.C., Blake D.J., Reid K.B.M.;
RT   "Characterization and organization of the genes encoding the A-, B-
RT   and C-chains of human complement subcomponent C1q. The complete
RT   derived amino acid sequence of human C1q.";
RL   Biochem. J. 274:481-490(1991).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Cerebellum;
RA   Nishi T., Nakagawa S., Senoh A., Mizuguchi H., Inagaki H., Suzuki Y.,
RA   Hata H., Nakagawa K., Mizuno S., Morinaga M., Kawamura M.,
RA   Sugiyama T., Irie R., Otsuki T., Sato H., Nishikawa T., Sugiyama A.,
RA   Kawakami B., Nagai K., Isogai T., Sugano S.;
RT   "NEDO human cDNA sequencing project.";
RL   Submitted (OCT-2001) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain;
RA   Strausberg R.;
RL   Submitted (MAY-2001) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE OF 29-122.
RX   MEDLINE=80020137; PubMed=486087;
RA   Reid K.B.M.;
RT   "Complete amino acid sequences of the three collagen-like regions
RT   present in subcomponent C1q of the first component of human
RT   complement.";
RL   Biochem. J. 179:367-371(1979).
RN   [5]
RP   REVIEW OF C1Q DEFICIENCY.
RX   MEDLINE=98450587; PubMed=9777412;
RA   Petry F.;
RT   "Molecular basis of hereditary C1q deficiency.";
RL   Immunobiology 199:286-294(1998).
CC   -!- FUNCTION: C1Q ASSOCIATES WITH THE PROENZYMES C1R AND C1S TO YIELD
CC       C1, THE FIRST COMPONENT OF THE SERUM COMPLEMENT SYSTEM. THE
CC       COLLAGEN-LIKE REGIONS OF C1Q INTERACT WITH THE CA(2+)-DEPENDENT
CC       C1R(2)C1S(2) PROENZYME COMPLEX, AND EFFICIENT ACTIVATION OF C1
CC       TAKES PLACE ON INTERACTION OF THE GLOBULAR HEADS OF C1Q WITH THE
CC       FC REGIONS OF IGG OR IGM ANTIBODY PRESENT IN IMMUNE COMPLEXES.
CC   -!- SUBUNIT: C1 IS A CALCIUM-DEPENDENT TRIMOLECULAR COMPLEX OF C1Q, R
CC       AND S IN THE MOLAR RATION OF 1:2:2. C1Q SUBCOMPONENT IS COMPOSED
CC       OF NINE SUBUNITS, SIX OF WHICH ARE DISULFIDE-LINKED DIMERS OF THE
CC       A AND B CHAINS, AND THREE OF WHICH ARE DISULFIDE-LINKED DIMERS OF
CC       THE C CHAIN.
CC   -!- PTM: O-LINKED GLYCANS CONSIST OF GLC-GAL DISACCHARIDES.
CC   -!- DISEASE: DEFECTS IN C1QG ARE A CAUSE OF C1Q DEFICIENCY. IT IS A
CC       RARE GENETIC DISORDER WHICH IS ASSOCIATED WITH RECURRENT
CC       INFECTIONS AND A HIGH PREVALENCE OF LUPUS ERYTHEMATOSUS-LIKE
CC       SYMPTOMS. IT IS CHARACTERIZED BY A LOSS OF ACTIVATION OF THE
CC       COMPLEMENT CLASSICAL PATHWAY.
CC   -!- SIMILARITY: CONTAINS 1 COLLAGENOUS DOMAIN.
CC   -!- SIMILARITY: CONTAINS 1 C1Q DOMAIN.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AK057792; BAB71575.1; -.
DR   EMBL; BC009016; AAH09016.1; -.
DR   PIR; A03207; C1HUQC.
DR   PIR; S14351; S14351.
DR   Genew; HGNC:1245; C1QG.
DR   MIM; 120575; -.
DR   InterPro; IPR001073; C1q.
DR   InterPro; IPR000087; Collagen.
DR   Pfam; PF00386; C1q; 1.
DR   Pfam; PF01391; Collagen; 1.
DR   PRINTS; PR00007; COMPLEMNTC1Q.
DR   SMART; SM00110; C1Q; 1.
DR   PROSITE; PS01113; C1Q; 1.
KW   Complement pathway; Plasma; Hydroxylation; Glycoprotein; Collagen;
KW   Repeat; Signal; Disease mutation.
FT   SIGNAL        1     28
FT   CHAIN        29    245       COMPLEMENT C1Q SUBCOMPONENT, C CHAIN.
FT   DOMAIN       31    112       COLLAGEN-LIKE.
FT   DOMAIN      113    245       C1Q.
FT   DISULFID     32     32       INTERCHAIN (WITH OTHER C CHAIN).
FT   MOD_RES      36     36       HYDROXYLATION.
FT   MOD_RES      39     39       HYDROXYLATION.
FT   MOD_RES      42     42       HYDROXYLATION.
FT   MOD_RES      45     45       HYDROXYLATION.
FT   MOD_RES      54     54       HYDROXYLATION.
FT   MOD_RES      57     57       HYDROXYLATION.
FT   MOD_RES      63     63       HYDROXYLATION.
FT   MOD_RES      66     66       HYDROXYLATION.
FT   MOD_RES      71     71       HYDROXYLATION.
FT   MOD_RES      75     75       HYDROXYLATION.
FT   CARBOHYD     75     75       O-LINKED (GAL. . .).
FT   MOD_RES      81     81       HYDROXYLATION.
FT   MOD_RES      84     84       HYDROXYLATION.
FT   CARBOHYD     84     84       O-LINKED (GAL. . .).
FT   MOD_RES      93     93       HYDROXYLATION.
FT   MOD_RES      96     96       HYDROXYLATION.
FT   MOD_RES      99     99       HYDROXYLATION.
FT   MOD_RES     105    105       HYDROXYLATION.
FT   VARIANT      43     43       G -> R (IN C1Q DEFICIENCY).
FT                                /FTId=VAR_008542.
FT   CONFLICT     14     14       K -> R (IN REF. 2).
FT   CONFLICT     23     23       P -> A (IN REF. 1).
FT   CONFLICT     57     57       K -> P (IN REF. 4).
FT   CONFLICT     66     66       P -> K (IN REF. 4).
FT   CONFLICT     72     72       K -> P (IN REF. 4).
FT   CONFLICT     84     84       P -> K (IN REF. 4).
FT   CONFLICT     87     87       N -> D (IN REF. 4).
FT   CONFLICT     90     90       M -> N (IN REF. 4).
FT   CONFLICT    215    215       E -> G (IN REF. 2).
SQ   SEQUENCE   245 AA;  25774 MW;  FA17117EB7ABFC12 CRC64;

  Query Match             30.0%;  Score 397;  DB 1;  Length 245;
  Best Local Similarity   39.8%;  Pred. No. 4e-21;
  Matches  100;  Conservative   33;  Mismatches   84;  Indels   34;  Gaps   10;

Qy          1 MRPLLVLLLLGLAAGSPPLDDNKIPSLC---PGHPGLPGTPGHHGSQGLPGRDGRDGRDG 57
              :: ||:|||| |          :  : |   || ||||| ||  |  ||||  |  |
```

```
Db          13 LKLLLLLLLLPLR--------GQANTGCYGIPGMPGLPGAPGKDGYDGLPGPKGEPGIPA 64

Qy          58 APGAPGEKGEGGRPGLPGPRGDPGPRGEAGPAGPTGPAGECSVPP---------RSAFSA 108

Db          65 IPGIRGPKGQKGEPGLPGHPGKNGPMGPPGMPGVPGPMGIPGEPGEEGRYKQKFQSVFTV 124

Qy         109 KRSESRVPPPSDAPLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDL 168

Db         125 TR-QTHQPPAPNSLIRFNAVLTNPQGDYDTSTGKFTCKVPGLYYFVYHAS-HTANLCVLL 182

Qy         169 VKNGESIASFFQFFGGWPKPASL-SGGAMVRLEPEDQVWVQVGVGDY---IGIYASIKTD 224

Db         183 YRSGVKVVT---FCGHTSKTNQVNSGGVLLRLQVGEEVW--LAVNDYYDMVGIQGS---D 234

Qy         225 STFSGFLVYSD 235

Db         235 SVFSGFLLFPD 245


RESULT 12
CA18_RABIT
ID   CA18_RABIT     STANDARD;      PRT;   744 AA.
AC   P14282;
DT   01-JAN-1990 (Rel. 13, Created)
DT   01-JAN-1990 (Rel. 13, Last sequence update)
DT   01-NOV-1995 (Rel. 32, Last annotation update)
DE   Collagen alpha 1(VIII) chain precursor (Endothelial collagen).
GN   COL8A1.
OS   Oryctolagus cuniculus (Rabbit).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX   NCBI_TaxID=9986;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=89380199; PubMed=2476437;
RA   Yamaguchi N., Benya P.D., van der Rest M., Ninomiya Y.;
RT   "The cloning and sequencing of alpha 1(VIII) collagen cDNAs
RT   demonstrate that type VIII collagen is a short chain collagen and
RT   contains triple-helical and carboxyl-terminal non-triple-helical
RT   domains similar to those of type X collagen.";
RL   J. Biol. Chem. 264:16022-16029(1989).
CC   -!- FUNCTION: MAJOR COMPONENT OF THE DESCEMET'S MEMBRANE (BASEMENT
CC       MEMBRANE) OF CORNEAL ENDOTHELIAL CELLS.
CC   -!- SUBUNIT: MAY FORM HOMOTRIMERS, OR HETEROTRIMERS IN ASSOCIATION
CC       WITH ALPHA 2(VIII) TYPE COLLAGENS.
CC   -!- PTM: PROLINES AT THE THIRD POSITION OF THE TRIPEPTIDE REPEATING
CC       UNIT (G-X-Y) ARE HYDROXYLATED IN SOME OR ALL OF THE CHAINS.
CC   -!- MISCELLANEOUS: FOUR CONSECUTIVE GLY-PRO-PRO TRIPLETS ARE PRESENT
CC       AT THE C-TERMINUS OF THE TRIPLE-HELICAL REGION. THESE MAY PROVIDE
CC       HIGH THERMAL STABILITY OF THIS REGION OF THIS REGION.
CC   -!- SIMILARITY: STRONG, TO ALPHA 2 TYPES VIII AND X COLLAGENS.
CC   -!- SIMILARITY: CONTAINS 1 C1Q DOMAIN.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; J05042; AAA31204.1; -.
DR   PIR; A34246; A34246.
DR   InterPro; IPR001073; C1q.
DR   InterPro; IPR000087; Collagen.
DR   Pfam; PF00386; C1q; 1.
DR   Pfam; PF01391; Collagen; 8.
DR   PRINTS; PR00007; COMPLEMNTC1Q.
DR   SMART; SM00110; C1Q; 1.
DR   PROSITE; PS01113; C1Q; 1.
KW   Extracellular matrix; Connective tissue; Repeat; Hydroxylation;
KW   Glycoprotein; Cell adhesion; Collagen; Signal.
FT   SIGNAL        1     20
FT   CHAIN         21    744        COLLAGEN ALPHA 1(VIII) CHAIN.
FT   DOMAIN        29    117        NONHELICAL REGION (NC2).
FT   DOMAIN        118   571        TRIPLE-HELICAL REGION.
FT   DOMAIN        572   744        NONHELICAL REGION (NC1).
FT   DOMAIN        609   744        C1Q.
SQ   SEQUENCE   744 AA;  73358 MW;  2A8CEF1EF8274E99 CRC64;

  Query Match             28.2%;  Score 374;  DB 1;  Length 744;
  Best Local Similarity   34.4%;  Pred. No. 4.7e-19;
  Matches   93;  Conservative   36;  Mismatches   79;  Indels   62;  Gaps    8;

Qy          24 IPSLC--PGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPGAPGEKGEGGRPGLP------- 74

Db         475 VPGLLGPKGEPGIPGDQGLQGPPGIPGITGPSGPIGPPGIPGPKGEPGLPGPPGFPGVGK 534

Qy          75 -------GPRGDP---GPRGE---AGPAGPTGPAGECSVPPRS---------------- 104

Db         535 PGVAGLHGPPGKPGALGPQGQPGLPGPPGPPGPPGPPAVMPPTPAPQGEYLPDMGLGIDG 594

Qy         105 -----AFSAKRSESRVP--------------PPSDAPLPFDRVLVNEQGHYDAVTGKFT 144

Db         595 VKTPHAYAAKKGKNGGPAYEMPAFTAELTAPFPPVGAPIKFDRLLYNGRQNYNPQTGIFT 654

Qy         145 CQVPGVYYFAVHATVYRASLQFDLVKNGESIA-SFFQFFGGWPKPASLSGGAMVRLEPED 203

Db         655 CEVPGVYYFAYHVHCKGGNVWVALFKNNEPVMYTYDEYKKGFLDQA--SGSAVLLLRPGD 712

Qy         204 QVWVQVGVGDYIGIYASIKTDSTFSGFLVY 233

Db         713 RVFLQMPSEQAAGLYAGQYVHSSFSGYLLY 742


RESULT 13
COLE_LEPMA
ID   COLE_LEPMA     STANDARD;      PRT;   419 AA.
AC   P98085; Q91080;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   15-JUL-1998 (Rel. 36, Last annotation update)
DE   Inner ear-specific collagen precursor (Saccular collagen).
OS   Lepomis macrochirus (Bluegill).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC   Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Percoidei;
OC   Centrarchidae; Lepomis.
OX   NCBI_TaxID=13106;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=95167486; PubMed=7863331;
RA   Davis J.G., Oberholtzer J.C., Burns F.R., Greene M.I.;
RT   "Molecular cloning and characterization of an inner ear-specific
RT   structural protein.";
RL   Science 267:1031-1034(1995).
RN   [2]
RP   CONCEPTUAL TRANSLATION.
RA   Gibson T.;
RL   Submitted (MAR-1995) to the SWISS-PROT data bank.
CC   -!- FUNCTION: FORMS A MICROSTRUCTURAL MATRIX WITHIN THE OTOLITHIC
CC       MEMBRANE (PROBABLY).
CC   -!- TISSUE SPECIFICITY: SPECIALIZED SECRETORY SUPPORTING CELLS AT THE
CC       OUTER PERIMETER OF THE SACCULAR EPITHELIUM.
CC   -!- SIMILARITY: CONTAINS 1 C1Q DOMAIN.
CC   -!- CAUTION: THIS IS A CONCEPTUAL TRANSLATION; A FRAMESHIFT HAD TO BE
CC       INTRODUCED FOR POSITIONS 391-419 SO AS TO MAXIMIZE THE SIMILARITY
CC       WITH OTHER SHORT-CHAIN COLLAGENS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
```

```
CC   ----------------------------------------------------------------------
DR   EMBL; U17431; AAA69978.1; ALT_FRAME.
DR   InterPro; IPR001073; C1q.
DR   InterPro; IPR000087; Collagen.
DR   Pfam; PF00386; C1q; 1.
DR   Pfam; PF01391; Collagen; 3.
DR   PRINTS; PR00007; COMPLEMNTC1Q.
DR   SMART; SM00110; C1Q; 1.
DR   PROSITE; PS01113; C1Q; 1.
KW   Extracellular matrix; Repeat; Collagen; Signal.
FT   SIGNAL        1     19       POTENTIAL.
FT   CHAIN        20    419       INNER EAR-SPECIFIC COLLAGEN.
FT   DOMAIN       20     57       NONHELICAL REGION (NC2).
FT   DOMAIN       58    274       TRIPLE-HELICAL REGION (COL1).
FT   DOMAIN      275    419       NONHELICAL REGION (NC1).
FT   DOMAIN      272    419       C1Q.
FT   CARBOHYD     37     37       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    320    320       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   419 AA;  43634 MW;  570CDB9675FC0F39 CRC64;

  Query Match             28.2%;  Score 373;  DB 1;  Length 419;
  Best Local Similarity   37.5%;  Pred. No. 3.1e-19;
  Matches   96;  Conservative   23;  Mismatches   99;  Indels   38;  Gaps    6;

Qy         15 GSPPLDDNK--IPSLCPGHPGLPGTPGHHGSQGLPGR--DGRDGRDGAPGAPGEKGEGGR 70
              | | |:  |  |    |  ||| || |  | ||| |    |  |  | ||  || |  |
Db        157 GEPGLNGTKGSIGREGPMGPGLAGTKGLKGEQGLKGECLQGEKGERGPPGLRGEMGLNGT 216

Qy         71 PGLPGPRGDPGP----------------------------RGEAGPAGPTGPAGECSV 100
                |: | ||:|||                           :|  || || || ||
Db        217 DGVKGERGEPGPLGGKGDTGARGPPGPPGGRGMAGLRGEKGLKGVRGPRGPKGPPGESVE 276

Qy        101 PPRSAFSAKRSESRVPPPSDAPLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVY 160
                |||||     ||  ||    |: ||:|  | :||:|    ||    |||| |: | ||
Db        277 QIRSAFSVGLFPSRSFPPPSLPVKFDKVFYNGEGHWDPTLNKFNVTYPGVYLFSYHITVR 336

Qy        161 RASLQFDLVKNG-ESIASFFQFFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYA 219
                ::  || ||   : :    :|     | |  |:: |   ||||::  : |: | |:
Db        337 NRPVRAALVVNGVRKLRTRDSLYGQDIDQA--SNLALLHLTDGDQVWLET-LRDWNGXYS 393

Qy        220 SIKTDSTFSGFLVYSD 235
                | : ||||||||:| |
Db        394 SSEDDSTFSGFLLYPD 409


RESULT 14
CA18_HUMAN
ID   CA18_HUMAN     STANDARD;      PRT;   744 AA.
AC   P27658; Q96D07;
DT   01-AUG-1992 (Rel. 23, Created)
DT   15-JUN-2002 (Rel. 41, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Collagen alpha 1(VIII) chain precursor (Endothelial collagen).
GN   COL8A1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91231001; PubMed=2029894;
RA   Muragaki Y., Mattei M.-G., Yamaguchi N., Olsen B.R., Ninomiya Y.;
RT   "The complete primary structure of the human alpha 1 (VIII) chain and
RT   assignment of its gene (COL8A1) to chromosome 3.";
RL   Eur. J. Biochem. 197:615-622(1991).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Lung;
RA   Strausberg R.;
RL   Submitted (SEP-2001) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: MAJOR COMPONENT OF THE DESCEMET'S MEMBRANE (BASEMENT
CC       MEMBRANE) OF CORNEAL ENDOTHELIAL CELLS.
CC   -!- SUBUNIT: MAY FORM HOMOTRIMERS, OR HETEROTRIMERS IN ASSOCIATION
CC       WITH ALPHA 2(VIII) TYPE COLLAGENS.
CC   -!- PTM: PROLINES AT THE THIRD POSITION OF THE TRIPEPTIDE REPEATING
CC       UNIT (G-X-Y) ARE HYDROXYLATED IN SOME OR ALL OF THE CHAINS.
CC   -!- MISCELLANEOUS: FOUR CONSECUTIVE GLY-PRO-PRO TRIPLETS ARE PRESENT
CC       AT THE C-TERMINUS OF THE TRIPLE-HELICAL REGION. THESE MAY PROVIDE
CC       THE HIGH THERMAL STABILITY OF THIS REGION.
CC   -!- SIMILARITY: STRONG, TO ALPHA 2 TYPES VIII AND X COLLAGENS.
CC   -!- SIMILARITY: CONTAINS 1 C1Q DOMAIN.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; X57527; CAA40748.1; -.
DR   EMBL; BC013581; AAH13581.1; -.
DR   PIR; S15435; S15435.
DR   Genew; HGNC:2215; COL8A1.
DR   MIM; 120251; -.
DR   InterPro; IPR001073; C1q.
DR   InterPro; IPR000087; Collagen.
DR   Pfam; PF00386; C1q; 1.
DR   Pfam; PF01391; Collagen; 8.
DR   PRINTS; PR00007; COMPLEMNTC1Q.
DR   ProDom; PD000007; Collagen; 1.
DR   SMART; SM00110; C1Q; 1.
DR   PROSITE; PS01113; C1Q; 1.
KW   Extracellular matrix; Connective tissue; Repeat; Hydroxylation;
KW   Glycoprotein; Cell adhesion; Collagen; Signal.
FT   SIGNAL        1     28
FT   CHAIN        29    744       COLLAGEN ALPHA 1(VIII) CHAIN.
FT   DOMAIN       29    117       NONHELICAL REGION (NC2).
FT   DOMAIN      118    571       TRIPLE-HELICAL REGION (COL1).
FT   DOMAIN      572    744       NONHELICAL REGION (NC1).
FT   DOMAIN      609    744       C1Q.
FT   CONFLICT    262    262       P -> L (IN REF. 1).
FT   CONFLICT    297    297       P -> R (IN REF. 1).
FT   CONFLICT    344    344       P -> A (IN REF. 1).
FT   CONFLICT    382    382       A -> S (IN REF. 1).
FT   CONFLICT    388    388       P -> S (IN REF. 1).
FT   CONFLICT    454    454       L -> F (IN REF. 1).
FT   CONFLICT    464    464       A -> H (IN REF. 1).
FT   CONFLICT    601    601       Y -> T (IN REF. 1).
FT   CONFLICT    631    631       A -> G (IN REF. 1).
SQ   SEQUENCE   744 AA;  73364 MW;  2BC1B0955DE2C9A3 CRC64;

  Query Match             27.8%;  Score 368;  DB 1;  Length 744;
  Best Local Similarity   33.3%;  Pred. No. 1.2e-18;
  Matches   90;  Conservative   34;  Mismatches   84;  Indels   62;  Gaps    6;

Qy         24 IPSLC--PGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPGAPGEKGEGGRPGLP------- 74
              :| |    | ||:||  |  |  |:||  |  |  | || || ||| | || |
Db        475 VPGLLGPKGEPGIPGDQGLQGPPGIPGIGGPSGPIGPPGIPGPKGEPGLPGPPGFPGIGK 534

Qy         75 ---------------GPRGDPGPRGEAGPAGPTGP------------------AGECS 99
                            |:| ||  |  || || ||                    |
Db        535 PGVAGLHGPPGKPGALGPQGQPGLPGPPGPPGPPGPPAVMPPTPPPQGEYLPDMGLGIDG 594

Qy        100 VPPRSAFSAKRSESRVP--------------PPSDAPLPFDRVLVNEQGHYDAVTGKFT 144
              | |  |: ||: ::  |               ||  ||: |:::| | : :|:  || ||
Db        595 VKPPHAYGAKKGKNGGPAYEMPAFTAELTAPFPPVGAPVKFNKLLYNGRQNYNPQTGIFT 654

Qy        145 CQVPGVYYFAVHATVYRASLQFDLVKNGESIA-SFFQFFGGWPKPASLSGGAMVRLEPED 203
              |:|||||||| |       ::   | || | :  :: ::  |:   |  || |:: | | |
Db        655 CEVPGVYYFAYHVHCKGGNVWVALFKNNEPVMYTYDEYKKGFLDQA--SGSAVLLLRPGD 712

Qy        204 QVWVQVGVGDYIGIYASIKTDSTFSGFLVY 233
              :|::|:      |:||     |:|||:|:|
Db        713 RVFLQMPSEQAAGLYAGQYVHSSFSGYLLY 742
```

```
RESULT 15
CA18_MOUSE
ID   CA18_MOUSE     STANDARD;      PRT;   743 AA.
AC   Q00780; Q9D2V4;
DT   01-NOV-1995 (Rel. 32, Created)
DT   15-JUN-2002 (Rel. 41, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Collagen alpha 1(VIII) chain precursor.
GN   COL8A1.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BALB/c;
RX   MEDLINE=92362626; PubMed=1499564;
RA   Muragaki Y., Shiota C., Inoue M., Ooshima A., Olsen B.R.,
RA   Ninomiya Y.;
RT   "Alpha 1(VIII)-collagen gene transcripts encode a short-chain
RT   collagen polypeptide and are expressed by various epithelial,
RT   endothelial and mesenchymal cells in newborn mouse tissues.";
RL   Eur. J. Biochem. 207:895-902(1992).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Kidney;
RX   MEDLINE=21085660; PubMed=11217851;
RA   Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA   Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA   Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA   Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA   Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA   Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA   Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA   Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA   Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA   Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA   Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA   Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA   Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA   Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA   Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA   Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA   Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA   Hayashizaki Y.;
RT   "Functional annotation of a full-length mouse cDNA collection.";
RL   Nature 409:685-690(2001).
CC   -!- FUNCTION: MAJOR COMPONENT OF THE DESCEMET'S MEMBRANE (BASEMENT
CC       MEMBRANE) OF CORNEAL ENDOTHELIAL CELLS.
CC   -!- SUBUNIT: MAY FORM HOMOTRIMERS, OR HETEROTRIMERS IN ASSOCIATION
CC       WITH ALPHA 2(VIII) TYPE COLLAGENS.
CC   -!- TISSUE SPECIFICITY: HIGH LEVELS IN CALVARIUM, EYE & SKIN OF
CC       NEWBORN MICE; ALSO IN VARIOUS EPITHELIAL, ENDOTHELIAL AND
CC       MESENCHYMAL CELLS.
CC   -!- PTM: PROLINES AT THE THIRD POSITION OF THE TRIPEPTIDE REPEATING
CC       UNIT (G-X-Y) ARE HYDROXYLATED IN SOME OR ALL OF THE CHAINS.
CC   -!- SIMILARITY: STRONG, TO ALPHA 2 TYPES VIII AND X COLLAGENS.
CC   -!- SIMILARITY: CONTAINS 1 C1Q DOMAIN.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X66976; CAA47387.1; -.
DR   EMBL; X66977; CAA47387.1; JOINED.
DR   EMBL; AK018742; BAB31383.1; -.
DR   PIR; S23779; S23779.
DR   MGD; MGI:88463; Col8a1.
DR   InterPro; IPR001073; C1q.
DR   InterPro; IPR000087; Collagen.
DR   Pfam; PF00386; C1q; 1.
DR   Pfam; PF01391; Collagen; 7.
DR   PRINTS; PR00007; COMPLEMNTC1Q.
DR   SMART; SM00110; C1Q; 1.
DR   PROSITE; PS01113; C1Q; 1.
KW   Extracellular matrix; Connective tissue; Repeat; Hydroxylation;
KW   Glycoprotein; Cell adhesion; Collagen; Signal.
FT   SIGNAL        1     28       BY SIMILARITY.
FT   CHAIN        29    743       COLLAGEN ALPHA 1(VIII) CHAIN.
FT   DOMAIN       29    117       NONHELICAL REGION (NC2).
FT   DOMAIN      118    571       TRIPLE-HELICAL REGION (COL1).
FT   DOMAIN      572    743       NONHELICAL REGION (NC1).
FT   DOMAIN      608    743       C1Q.
FT   CONFLICT      6      6       G -> R (IN REF. 2).
FT   CONFLICT     85     85       H -> Y (IN REF. 1).
FT   CONFLICT    109    109       K -> KG (IN REF. 2).
FT   CONFLICT    248    248       P -> L (IN REF. 1).
FT   CONFLICT    313    313       P -> A (IN REF. 2).
FT   CONFLICT    323    324       IP -> SR (IN REF. 1).
FT   CONFLICT    361    361       D -> H (IN REF. 1).
FT   CONFLICT    596    596       T -> P (IN REF. 2).
FT   CONFLICT    717    719       MPS -> NPF (IN REF. 1).
SQ   SEQUENCE   743 AA;  73454 MW;  F584D85BD53897F4 CRC64;

  Query Match             27.7%;  Score 366.5;  DB 1;  Length 743;
  Best Local Similarity   32.8%;  Pred. No. 1.5e-18;
  Matches   90;  Conservative   36;  Mismatches   77;  Indels   71;  Gaps    7;

Qy         29 PGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPGAPGEKGEGGRPGLPGPRGDP-------- 80
              || ||:||  |  |  |:||  |  |  | || |  |  | ||:|||:|:|
Db        470 PGLPGVPGLLGPKGEPGIPGDQGLQGPPGIPGIVGPSGPIGPPGIPGPKGEPGLPGPPGF 529

Qy         81 ------------------F--GPRGE---AGPAGPTGPAGECSVPPR------------ 103
                                  ||:|:    || || || |  :| |
Db        530 PGVGKPGVAGLHGPPGKPGALGPQGQPGLPGPPGPPGPPGPPAVMPTPSPQGEYLPDMGL 589

Qy        104 --------SAFSAKR-------------SESRVP-PPSDAPLPFDRVLVNEQGHYDAVT 140
                    |:: |:              :|  || ||  ||: ||::| | : :|:  |
Db        590 GIDGVKTPHAYAGKKGKHGGPAYEMPAFTAELTVPFPPVGAPVKFDKLLYNGRQNYNPQT 649

Qy        141 GKFTCQVPGVYYFAVHATVYRASLQFDLVKNGESIA-SFFQFFGGWPKPASLSGGAMVRL 199
              | |||:|||||||| |      ::   | || | :  :: ::  |:   |  || |:: |
Db        650 GIFTCEVPGVYYFAYHVHCKGGNVWVALFKNNEPMMYTYDEYKKGFLDQA--SGSAVLLL 707

Qy        200 EPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVY 233
               | |||::|:      |:||     |:|||:|:|
Db        708 RPGDQVFLQMPSEQAAGLYAGQYVHSSFSGYLLY 741


Search completed: June 18, 2003, 14:53:00
Job time : 24 secs
```

GenCore version 5.1.6


OM protein - protein search, using sw model

Run on: June 18, 2003, 14:49:37 ; Search time 26 Seconds
(without alignments)
274.991 Million cell updates/sec

Title: US-09-943-851A-42
Perfect score: 1325
Sequence: 1 MRPLLVLLLLGLAAGSPPLD.....DSTFSGFLVYSDWHSSPVFA 243

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 1325 | 100.0 | 243 | 4 | US-09-140-804-2 | Sequence 2, Appli |
| 2 | 1325 | 100.0 | 243 | 4 | US-09-336-536-3 | Sequence 3, Appli |
| 3 | 1271 | 95.9 | 243 | 4 | US-09-188-930-295 | Sequence 295, App |
| 4 | 1258 | 94.9 | 243 | 4 | US-09-336-536-10 | Sequence 10, Appl |
| 5 | 1256 | 94.8 | 228 | 4 | US-09-336-536-4 | Sequence 4, Appli |
| 6 | 1200 | 90.6 | 228 | 4 | US-09-336-536-11 | Sequence 11, Appl |
| 7 | 675 | 50.9 | 128 | 4 | US-09-336-536-7 | Sequence 7, Appli |
| 8 | 650 | 49.1 | 128 | 4 | US-09-336-536-14 | Sequence 14, Appl |
| 9 | 433.5 | 32.7 | 247 | 2 | US-08-463-911-2 | Sequence 2, Appli |
| 10 | 428.5 | 32.3 | 247 | 4 | US-09-140-804-8 | Sequence 8, Appli |
| 11 | 428.5 | 32.3 | 247 | 4 | US-09-118-408-3 | Sequence 3, Appli |
| 12 | 428.5 | 32.3 | 247 | 4 | US-09-506-855-3 | Sequence 3, Appli |
| 13 | 424 | 32.0 | 244 | 2 | US-08-463-911-7 | Sequence 7, Appli |
| 14 | 424 | 32.0 | 244 | 4 | US-09-140-804-3 | Sequence 3, Appli |
| 15 | 424 | 32.0 | 244 | 4 | US-09-336-536-20 | Sequence 20, Appl |
| 16 | 424 | 32.0 | 244 | 4 | US-09-530-423-1 | Sequence 1, Appli |
| 17 | 417 | 31.5 | 231 | 4 | US-09-530-423-2 | Sequence 2, Appli |
| 18 | 412 | 31.1 | 246 | 2 | US-08-463-911-4 | Sequence 4, Appli |
| 19 | 399 | 30.1 | 294 | 4 | US-09-188-930-294 | Sequence 294, App |
| 20 | 347 | 26.2 | 60 | 4 | US-09-336-536-6 | Sequence 6, Appli |
| 21 | 324 | 24.5 | 60 | 4 | US-09-336-536-13 | Sequence 13, Appl |
| 22 | 316 | 23.8 | 423 | 1 | US-08-383-744-2 | Sequence 2, Appli |
| 23 | 316 | 23.8 | 423 | 2 | US-08-999-336-2 | Sequence 2, Appli |
| 24 | 316 | 23.8 | 423 | 5 | PCT-US96-01427-2 | Sequence 2, Appli |
| 25 | 314 | 23.7 | 245 | 4 | US-09-140-804-4 | Sequence 4, Appli |
| 26 | 280.5 | 21.2 | 215 | 4 | US-09-140-804-5 | Sequence 5, Appli |
| 27 | 277.5 | 20.9 | 185 | 2 | US-08-463-911-3 | Sequence 3, Appli |
| 28 | 276 | 20.8 | 236 | 4 | US-09-140-804-6 | Sequence 6, Appli |
| 29 | 266.5 | 20.1 | 198 | 4 | US-09-188-930-138 | Sequence 138, App |
| 30 | 257.5 | 19.4 | 130 | 4 | US-09-485-316A-13 | Sequence 13, Appl |
| 31 | 251.5 | 19.0 | 130 | 4 | US-09-485-316A-12 | Sequence 12, Appl |
| 32 | 243.5 | 18.4 | 130 | 4 | US-09-485-316A-11 | Sequence 11, Appl |
| 33 | 240 | 18.1 | 161 | 4 | US-09-415-551-3 | Sequence 3, Appli |
| 34 | 229.5 | 17.3 | 280 | 4 | US-09-247-155-178 | Sequence 178, App |
| 35 | 229.5 | 17.3 | 281 | 4 | US-09-118-408-2 | Sequence 2, Appli |
| 36 | 229.5 | 17.3 | 281 | 4 | US-09-506-855-2 | Sequence 2, Appli |
| 37 | 228 | 17.2 | 684 | 1 | US-08-555-669-12 | Sequence 12, Appl |
| 38 | 228 | 17.2 | 684 | 3 | US-09-073-663-12 | Sequence 12, Appl |
| 39 | 224 | 16.9 | 228 | 4 | US-09-219-849-38 | Sequence 38, Appl |
| 40 | 223 | 16.8 | 281 | 4 | US-09-118-408-44 | Sequence 44, Appl |
| 41 | 223 | 16.8 | 281 | 4 | US-09-506-855-44 | Sequence 44, Appl |
| 42 | 221 | 16.7 | 124 | 4 | US-09-485-316A-9 | Sequence 9, Appli |
| 43 | 219.5 | 16.6 | 1057 | 3 | US-08-931-820-4 | Sequence 4, Appli |
| 44 | 219.5 | 16.6 | 1078 | 3 | US-08-963-825-21 | Sequence 21, Appl |
| 45 | 219.5 | 16.6 | 1078 | 4 | US-09-500-811-21 | Sequence 21, Appl |

ALIGNMENTS

RESULT 1
US-09-140-804-2
; Sequence 2, Application US/09140804
; Patent No. 6197930
; GENERAL INFORMATION:
; APPLICANT: Sheppard, Paul O.
; APPLICANT: Humes, Jacqueline M.
; TITLE OF INVENTION: ADIPOCYTE-SPECIFIC PROTEIN HOMOLOGS
; FILE REFERENCE: 97-49
; CURRENT APPLICATION NUMBER: US/09/140,804
; CURRENT FILING DATE: 1998-08-26
; EARLIER APPLICATION NUMBER: 60/056,983
; EARLIER FILING DATE: 1997-08-26
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 243
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-140-804-2

Query Match 100.0%; Score 1325; DB 4; Length 243;
Best Local Similarity 100.0%; Pred. No. 9.5e-114;
Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MRPLLVLLLLGLAAGSPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MRPLLVLLLLGLAAGSPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPG 60

Qy         61 APGEKGEGGRPGLPGPRGDPGPRGEAGPAGPTGPAGECSVPPRSAFSAKRSESRVPPPSD 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 APGEKGEGGRPGLPGPRGDPGPRGEAGPAGPTGPAGECSVPPRSAFSAKRSESRVPPPSD 120

Qy        121 APLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGESIASFFQ 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 APLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGESIASFFQ 180

Qy        181 FFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWHSSP 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 FFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWHSSP 240

Qy        241 VFA 243
              |||
Db        241 VFA 243
```

RESULT 2
US-09-336-536-3
; Sequence 3, Application US/09336536
; Patent No. 6406884

```
; GENERAL INFORMATION:
;  APPLICANT: Leiby, K.
;  APPLICANT: McKay, C.
;  APPLICANT: Bossone, S.
;  TITLE OF INVENTION: SECRETED PROTEINS AND USES THEREOF
;  FILE REFERENCE: 7853-144
;  CURRENT APPLICATION NUMBER: US/09/336,536
;  CURRENT FILING DATE: 1999-06-18
;  NUMBER OF SEQ ID NOS: 75
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
;   LENGTH: 243
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-336-536-3

  Query Match             100.0%;  Score 1325;  DB 4;  Length 243;
  Best Local Similarity   100.0%;  Pred. No. 9.5e-114;
  Matches  243;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MRPLLVLLLLGLAAGSPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MRPLLVLLLLGLAAGSPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPG 60

Qy         61 APGEKGEGGRPGLPGPRGDPGPRGEAGPAGPTGPAGECSVPPRSAFSAKRSESRVPPPSD 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 APGEKGEGGRPGLPGPRGDPGPRGEAGPAGPTGPAGECSVPPRSAFSAKRSESRVPPPSD 120

Qy        121 APLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGESIASFFQ 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 APLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGESIASFFQ 180

Qy        181 FFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWHSSP 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 FFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWHSSP 240

Qy        241 VFA 243
              |||
Db        241 VFA 243


RESULT 3
US-09-188-930-295
; Sequence 295, Application US/09188930A
; Patent No. 6150502
; GENERAL INFORMATION:
;  APPLICANT: Watson, James D.
;  APPLICANT: Strachan, Lorna
;  APPLICANT: Sleeman, Matthew
;  APPLICANT: Onrust, Rene
;  APPLICANT: Murison, James Greg
;  TITLE OF INVENTION: Compositions Isolated From Skin Cells
;  TITLE OF INVENTION: and Methods For Their Use
;  FILE REFERENCE: 11000.1011c1
;  CURRENT APPLICATION NUMBER: US/09/188,930A
;  CURRENT FILING DATE: 1998-11-09
;  NUMBER OF SEQ ID NOS: 348
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 295
;   LENGTH: 243
;   TYPE: PRT
;   ORGANISM: Rat
US-09-188-930-295

  Query Match             95.9%;  Score 1271;  DB 4;  Length 243;
  Best Local Similarity   95.1%;  Pred. No. 8.2e-109;
  Matches  231;  Conservative    5;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          1 MRPLLVLLLLGLAAGSPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPG 60
              ||||| |||||||:|||||||||||||||| ||||||||||||||||||||||||||||||
Db          1 MRPLLALLLLGLASGSPPLDDNKIPSLCPGQPGLPGTPGHHGSQGLPGRDGRDGRDGAPG 60

Qy         61 APGEKGEGGRPGLPGPRGDPGPRGEAGPAGPTGPAGECSVPPRSAFSAKRSESRVPPPSD 120
              ||||||||||||||||||||:||||||| | ||||||||||||||||||||||||:|
Db         61 APGEKGEGGRPGLPGPRGEPGPRGEAGPVGAIGPAGECSVPPRSAFSAKRSESRVPPPAD 120

Qy        121 APLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGESIASFFQ 180
              ||||||||:|||||||| ||||||||||||||||||||||||||||||:|||||||
Db        121 TPLPFDRVLLNEQGHYDATTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGQSIASFFQ 180

Qy        181 FFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWHSSP 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 FFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWHSSP 240

Qy        241 VFA 243
              |||
Db        241 VFA 243


RESULT 4
US-09-336-536-10
; Sequence 10, Application US/09336536
; Patent No. 6406884
; GENERAL INFORMATION:
;  APPLICANT: Leiby, K.
;  APPLICANT: McKay, C.
;  APPLICANT: Bossone, S.
;  TITLE OF INVENTION: SECRETED PROTEINS AND USES THEREOF
;  FILE REFERENCE: 7853-144
;  CURRENT APPLICATION NUMBER: US/09/336,536
;  CURRENT FILING DATE: 1999-06-18
;  NUMBER OF SEQ ID NOS: 75
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10
;   LENGTH: 243
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-09-336-536-10

  Query Match             94.9%;  Score 1258;  DB 4;  Length 243;
  Best Local Similarity   93.8%;  Pred. No. 1.3e-107;
  Matches  228;  Conservative    6;  Mismatches    9;  Indels    0;  Gaps    0;

Qy          1 MRPLLVLLLLGLAAGSPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPG 60
              ||||| |||||| :|||||||||||||||| ||||||||||||||||||||||||||||||
Db          1 MRPLLALLLLGLVSGSPPLDDNKIPSLCPGQPGLPGTPGHHGSQGLPGRDGRDGRDGAPG 60

Qy         61 APGEKGEGGRPGLPGPRGDPGPRGEAGPAGPTGPAGECSVPPRSAFSAKRSESRVPPPSD 120
              ||||||||||||||||||:||||||| | ||||||||||||||||||||||||:|
Db         61 APGEKGEGGRPGLPGPRGEPGPRGEAGPMGAIGPAGECSVPPRSAFSAKRSESRVPPPAD 120

Qy        121 APLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGESIASFFQ 180
              ||||||||:|||||||| ||||||||||||||||||||||||||||||:|||||||
Db        121 TPLPFDRVLLNEQGHYDPTTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGQSIASFFQ 180

Qy        181 FFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWHSSP 240
              :||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 YFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWHSSP 240

Qy        241 VFA 243
              |||
Db        241 VFA 243


RESULT 5
US-09-336-536-4
; Sequence 4, Application US/09336536
; Patent No. 6406884
; GENERAL INFORMATION:
;  APPLICANT: Leiby, K.
;  APPLICANT: McKay, C.
;  APPLICANT: Bossone, S.
;  TITLE OF INVENTION: SECRETED PROTEINS AND USES THEREOF
;  FILE REFERENCE: 7853-144
;  CURRENT APPLICATION NUMBER: US/09/336,536
;  CURRENT FILING DATE: 1999-06-18
```

```
;  NUMBER OF SEQ ID NOS: 75
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
;   LENGTH: 228
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-336-536-4

  Query Match             94.8%;  Score 1256;  DB 4;  Length 228;
  Best Local Similarity   100.0%;  Pred. No. 1.8e-107;
  Matches  228;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         16 SPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPGAPGEKGEGGRPGLPG 75
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 SPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPGAPGEKGEGGRPGLPG 60

Qy         76 PRGDPGPRGEAGPAGPTGPAGECSVPPRSAFSAKRSESRVPPPSDAPLPFDRVLVNEQGH 135
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 PRGDPGPRGEAGPAGPTGPAGECSVPPRSAFSAKRSESRVPPPSDAPLPFDRVLVNEQGH 120

Qy        136 YDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGESIASFFQFFGGWPKPASLSGGA 195
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 YDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGESIASFFQFFGGWPKPASLSGGA 180

Qy        196 MVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWHSSPVFA 243
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 MVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWHSSPVFA 228


RESULT 6
US-09-336-536-11
; Sequence 11, Application US/09336536
; Patent No. 6406884
; GENERAL INFORMATION:
;  APPLICANT: Leiby, K.
;  APPLICANT: McKay, C.
;  APPLICANT: Bossone, S.
;  TITLE OF INVENTION: SECRETED PROTEINS AND USES THEREOF
;  FILE REFERENCE: 7853-144
;  CURRENT APPLICATION NUMBER: US/09/336,536
;  CURRENT FILING DATE: 1999-06-18
;  NUMBER OF SEQ ID NOS: 75
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
;   LENGTH: 228
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-09-336-536-11

  Query Match             90.6%;  Score 1200;  DB 4;  Length 228;
  Best Local Similarity   94.7%;  Pred. No. 2.4e-102;
  Matches  216;  Conservative    5;  Mismatches    7;  Indels    0;  Gaps    0;

Qy         16 SPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPGAPGEKGEGGRPGLPG 75
              |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db          1 SPPLDDNKIPSLCPGQPGLPGTPGHHGSQGLPGRDGRDGRDGAPGAPGEKGEGGRPGLPG 60

Qy         76 PRGDPGPRGEAGPAGPTGPAGECSVPPRSAFSAKRSESRVPPPSDAPLPFDRVLVNEQGH 135
              |||:|||||||||| |  |||||||||||||||||||||||:| ||||||||:|||||
Db         61 PRGEPGPRGEAGPMGAIGPAGECSVPPRSAFSAKRSESRVPPPADTPLPFDRVLLNEQGH 120

Qy        136 YDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGESIASFFQFFGGWPKPASLSGGA 195
              ||  |||||||||||||||||||||||||||:|||||||:|||||||||||||||||||
Db        121 YDPTTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGQSIASFFQYFGGWPKPASLSGGA 180

Qy        196 MVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWHSSPVFA 243
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 MVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWHSSPVFA 228


RESULT 7
US-09-336-536-7
; Sequence 7, Application US/09336536
; Patent No. 6406884
; GENERAL INFORMATION:
;  APPLICANT: Leiby, K.
;  APPLICANT: McKay, C.
;  APPLICANT: Bossone, S.
;  TITLE OF INVENTION: SECRETED PROTEINS AND USES THEREOF
;  FILE REFERENCE: 7853-144
;  CURRENT APPLICATION NUMBER: US/09/336,536
;  CURRENT FILING DATE: 1999-06-18
;  NUMBER OF SEQ ID NOS: 75
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
;   LENGTH: 128
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-336-536-7

  Query Match             50.9%;  Score 675;  DB 4;  Length 128;
  Best Local Similarity   100.0%;  Pred. No. 1.2e-54;
  Matches  128;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        105 AFSAKRSESRVPPPSDAPLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASL 164
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AFSAKRSESRVPPPSDAPLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASL 60

Qy        165 QFDLVKNGESIASFFQFFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTD 224
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 QFDLVKNGESIASFFQFFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTD 120

Qy        225 STFSGFLV 232
              ||||||||
Db        121 STFSGFLV 128


RESULT 8
US-09-336-536-14
; Sequence 14, Application US/09336536
; Patent No. 6406884
; GENERAL INFORMATION:
;  APPLICANT: Leiby, K.
;  APPLICANT: McKay, C.
;  APPLICANT: Bossone, S.
;  TITLE OF INVENTION: SECRETED PROTEINS AND USES THEREOF
;  FILE REFERENCE: 7853-144
;  CURRENT APPLICATION NUMBER: US/09/336,536
;  CURRENT FILING DATE: 1999-06-18
;  NUMBER OF SEQ ID NOS: 75
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14
;   LENGTH: 128
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-09-336-536-14

  Query Match             49.1%;  Score 650;  DB 4;  Length 128;
  Best Local Similarity   94.5%;  Pred. No. 2.3e-52;
  Matches  121;  Conservative    4;  Mismatches    3;  Indels    0;  Gaps    0;

Qy        105 AFSAKRSESRVPPPSDAPLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASL 164
              ||||||||||||||:| ||||||||:|||||| ||||||||||||||||||||||||
Db          1 AFSAKRSESRVPPPADTPLPFDRVLLNEQGHYDPTTGKFTCQVPGVYYFAVHATVYRASL 60

Qy        165 QFDLVKNGESIASFFQFFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTD 224
              ||||||||:|||||||:||||||||||||||||||||||||||||||||||||||||||
Db         61 QFDLVKNGQSIASFFQYFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTD 120

Qy        225 STFSGFLV 232
              ||||||||
Db        121 STFSGFLV 128


RESULT 9
US-08-463-911-2
```

```
; Sequence 2, Application US/08463911
; Patent No. 5869330
;  GENERAL INFORMATION:
;    APPLICANT:  Scherer, Philipp E.
;    APPLICANT:  Lodish, Harvey F.
;    TITLE OF INVENTION:  A NOVEL SERUM PROTEIN PRODUCED
;    TITLE OF INVENTION:  EXCLUSIVELY IN ADIPOCYTES
;    NUMBER OF SEQUENCES:  7
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Hamilton, Brook, Smith & Reynolds, P.C.
;      STREET:  Two Militia Drive
;      CITY:  Lexington
;      STATE:  Massachusetts
;      COUNTRY:  USA
;      ZIP:  02173
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/463,911
;      FILING DATE:
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Granahan, Patricia
;      REGISTRATION NUMBER:  32,227
;      REFERENCE/DOCKET NUMBER:  WHI95-05
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 861-6240
;      TELEFAX:  (617) 861-9540
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  247 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-463-911-2

  Query Match             32.7%;  Score 433.5;  DB 2;  Length 247;
  Best Local Similarity   40.6%;  Pred. No. 3.4e-32;
  Matches  102;  Conservative   32;  Mismatches   92;  Indels   25;  Gaps    8;

Qy          1 MRPLLVLLLL-GLAAGSPPLDDNKIPSLCPGHPG-----LPGTPGHHGSQGLPGRDGRDG 54
              :: || ||:|   |       :   |:| |   |     : | ||| |  | |||||||
Db          4 LQALLFLLILPSHAEDDVTTTEELAPALVPPPKGTCAGWMAGIPGHPGHNGTPGRDGRD- 62

Qy         55 RDGAPGAPGEKGEGGRPGLPGPRGD---------PGPRGEAGPAGPTGPAGECSVPPRSA 105
                | |||||| |  || ||:|:          ||||  |  |  |  || :   |||
Db         63 -----GTPGEKGEKGDAGLLGPKGETGDVGMTGAEGPRGFPGTPGRKGEPGEAAYMYRSA 117

Qy        106 FSAKRSESRVPPPSDAPLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQ 165
              ||    |:||  | : |: | ::  |:| |||  |||| | :||:|||: | |||   ::
Db        118 FSV-GLETRVTVP-NVPIRFTKIFYNQQNHYDGSTGKFYCNIPGLYYFSYHITVYMKDVK 175

Qy        166 FDLVKNGESIASFFQFFGGWPKPASLSGGAMVRLEPEDQVWVQV-GVGDYIGIYASIKTD 224
                | |  :::   :  :          ||  :: ||  |||||:|| | ||: |:||    |
Db        176 VSLFKKDKAVLFTYDQYQE-KNVDQASGSVLLHLEVGDQVWLQVYGDGDHNGLYADNVND 234

Qy        225 STFSGFLVYSD 235
              |||:|||:| |
Db        235 STFTGFLLYHD 245


RESULT 10
US-09-140-804-8
; Sequence 8, Application US/09140804
; Patent No. 6197930
; GENERAL INFORMATION:
;  APPLICANT: Sheppard, Paul O.
;  APPLICANT: Humes, Jacqueline M.
;  TITLE OF INVENTION: ADIPOCYTE-SPECIFIC PROTEIN HOMOLOGS
;  FILE REFERENCE: 97-49
;  CURRENT APPLICATION NUMBER: US/09/140,804
;  CURRENT FILING DATE: 1998-08-26
;  EARLIER APPLICATION NUMBER: 60/056,983
;  EARLIER FILING DATE: 1997-08-26
;  NUMBER OF SEQ ID NOS: 47
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 8
;   LENGTH: 247
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-140-804-8

  Query Match             32.3%;  Score 428.5;  DB 4;  Length 247;
  Best Local Similarity   40.2%;  Pred. No. 9.7e-32;
  Matches  101;  Conservative   32;  Mismatches   93;  Indels   25;  Gaps    8;

Qy          1 MRPLLVLLLL-GLAAGSPPLDDNKIPSLCPGHPG-----LPGTPGHHGSQGLPGRDGRDG 54
              :: || ||:|   |       :   |:| |   |     : | ||| |  | |||||||
Db          4 LQALLFLLILPSHAEDDVTTTEELAPALVPPPKGTCAGWMAGIPGHPGHNGTPGRDGRD- 62

Qy         55 RDGAPGAPGEKGEGGRPGLPGPRGDPGPRGEAGPAGP---------TGPAGECSVPPRSA 105
                | |||||| |  || ||:|:  |  |  |  ||           |  || :   |||
Db         63 -----GTPGEKGEKGDAGLLGPKGETGDVGMTGAEGPRGFPQTPGRKGEPGEAAYMYRSA 117

Qy        106 FSAKRSESRVPPPSDAPLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQ 165
              ||    |:||  | : |: | ::  |:| |||  |||| | :||:|||: | |||   ::
Db        118 FSV-GLETRVTVP-NVPIRFTKIFYNQQNHYDGSTGKFYCNIPGLYYFSYHITVYMKDVK 175


RESULT 11
US-09-118-408-3
; Sequence 3, Application US/09118408A
; Patent No. 6265544
; GENERAL INFORMATION:
;  APPLICANT: Sheppard, Paul O.
;  TITLE OF INVENTION: ADIPOCYTE-SPECIFIC PROTEIN HOMOLOGS
;  FILE REFERENCE: 97-30
;  CURRENT APPLICATION NUMBER: US/09/118,408A
;  CURRENT FILING DATE: 1998-07-17
;  EARLIER APPLICATION NUMBER: 60/053,154
;  EARLIER FILING DATE: 1997-07-18
;  NUMBER OF SEQ ID NOS: 47
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 3
;   LENGTH: 247
;   TYPE: PRT
;   ORGANISM: Homo sapien
US-09-118-408-3

  Query Match             32.3%;  Score 428.5;  DB 4;  Length 247;
  Best Local Similarity   40.2%;  Pred. No. 9.7e-32;
  Matches  101;  Conservative   32;  Mismatches   93;  Indels   25;  Gaps    8;

Qy          1 MRPLLVLLLL-GLAAGSPPLDDNKIPSLCPGHPG-----LPGTPGHHGSQGLPGRDGRDG 54
              :: || ||:|   |       :   |:| |   |     : | ||| |  | |||||||
Db          4 LQALLFLLILPSHAEDDVTTTEELAPALVPPPKGTCAGWMAGIPGHPGHNGTPGRDGRD- 62

Qy         55 RDGAPGAPGEKGEGGRPGLPGPRGDPGPRGEAGPAGP---------TGPAGECSVPPRSA 105
                | |||||| |  || ||:|:  |  |  |  ||           |  || :   |||
Db         63 -----GTPGEKGEKGDAGLLGPKGETGDVGMTGAEGPRGFPQTPGRKGEPGEAAYMYRSA 117

Qy        106 FSAKRSESRVPPPSDAPLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQ 165
              ||    |:||  | : |: | ::  |:| |||  |||| | :||:|||: | |||   ::
Db        118 FSV-GLETRVTVP-NVPIRFTKIFYNQQNHYDGSTGKFYCNIPGLYYFSYHITVYMKDVK 175
```

```
Qy        166 FDLVKNGESIASFFQFFGGWPKPASLSGGAMVRLEPEDQVWVQV-GVGDYIGIYASIKTD 224
              | |  :::   :  :          ||  :: ||  ||||:|| | ||: |:||    |
Db        176 VSLFKKDKAVLFTYDQYQE-KNVDQASGSVLLHLEVGDQVWLQVYGDGDHNGLYADNVND 234

Qy        225 STFSGFLVYSD 235
              |||:|||:| |
Db        235 STFTGFLLYHD 245


RESULT 12
US-09-506-855-3
; Sequence 3, Application US/09506855
; Patent No. 6448221
; GENERAL INFORMATION:
;  APPLICANT: Sheppard, Paul O.
;  APPLICANT: Lasser, Gerald W.
;  APPLICANT: Bishop, Paul D.
;  TITLE OF INVENTION: INHIBITORS FOR USE IN HEMOSTASIS AND
;  TITLE OF INVENTION: IMMUNE FUNCTION
;  FILE REFERENCE: 99-12
;  CURRENT APPLICATION NUMBER: US/09/506,855
;  CURRENT FILING DATE: 2000-02-17
;  NUMBER OF SEQ ID NOS: 50
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 3
;   LENGTH: 247
;   TYPE: PRT
;   ORGANISM: Homo sapien
US-09-506-855-3

  Query Match             32.3%;  Score 428.5;  DB 4;  Length 247;
  Best Local Similarity   40.2%;  Pred. No. 9.7e-32;
  Matches  101;  Conservative   32;  Mismatches   93;  Indels   25;  Gaps    8;

Qy          1 MRPLLVLLLL-GLAAGSPPLDDNKIPSLCPGHPG-----LPGTPGHHGSQGLPGRDGRDG 54
              :: || ||:|   |       :   |:| |   |     : | ||| |  | |||||||
Db          4 LQALLFLLILPSHAEDDVTTTEELAPALVPPPKGTCAGWMAGIPGHPGHNGTPGRDGRD- 62

Qy         55 RDGAPGAPGEKGEGGRPGLPGPRGDPGPRGEAGPAGP---------TGPAGECSVPPRSA 105
                | |||||| |  || ||:|: |  |  |  ||           |  || :   |||
Db         63 -----GTPGEKGEKGDAGLLGPKGETGDVGMTGAEGPRGFPQTPGRKGEPGEAAYMYRSA 117

Qy        106 FSAKRSESRVPPPSDAPLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQ 165
              ||    |:||  | : |: | ::  |:| |||  |||| | :||:|||: | |||   ::
Db        118 FSV-GLETRVTVP-NVPIRFTKIFYNQQNHYDGSTGKFYCNIPGLYYFSYHITVYMKDVK 175

Qy        166 FDLVKNGESIASFFQFFGGWPKPASLSGGAMVRLEPEDQVWVQV-GVGDYIGIYASIKTD 224
                | |  :::   :  :          ||  :: ||  ||||:|| | ||: |:||    |
Db        176 VSLFKKDKAVLFTYDQYQE-KNVDQASGSVLLHLEVGDQVWLQVYGDGDHNGLYADNVND 234

Qy        225 STFSGFLVYSD 235
              |||:|||:| |
Db        235 STFTGFLLYHD 245


RESULT 13
US-08-463-911-7
; Sequence 7, Application US/08463911
; Patent No. 5869330
;  GENERAL INFORMATION:
;    APPLICANT:  Scherer, Philipp E.
;    APPLICANT:  Lodish, Harvey F.
;    TITLE OF INVENTION:  A NOVEL SERUM PROTEIN PRODUCED
;    TITLE OF INVENTION:  EXCLUSIVELY IN ADIPOCYTES
;    NUMBER OF SEQUENCES:  7
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Hamilton, Brook, Smith & Reynolds, P.C.
;      STREET:  Two Militia Drive
;      CITY:  Lexington
;      STATE:  Massachusetts
;      COUNTRY:  USA
;      ZIP:  02173
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/463,911
;      FILING DATE:
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Granahan, Patricia
;      REGISTRATION NUMBER:  32,227
;      REFERENCE/DOCKET NUMBER:  WHI95-05
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 861-6240
;      TELEFAX:  (617) 861-9540
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  244 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-463-911-7

  Query Match             32.0%;  Score 424;  DB 2;  Length 244;
  Best Local Similarity   40.5%;  Pred. No. 2.5e-31;
  Matches  106;  Conservative   26;  Mismatches   72;  Indels   58;  Gaps   10;

Qy          6 VLLLLGLAAGSPPLDDNKIPSLCPGH--------PG-------------LPGTPGHHGSQ 44
              ||||| |                 |||        ||             : | ||| |
Db          7 VLLLLAL---------------PGHDQETTTQGPGVLLPLPKGACTGWMAGIPGHPGHN 50

Qy         45 GLPGRDGRDGRDGAPGAPGEKGEGGRPGLPGPRGD---------PGPRGEAGPAGPTGPA 95
              | |||||||      | ||||| | ||| ||:||          ||||  |  |  |
Db         51 GAPGRDGRD------GTPGEKGEKGDPGLIGPKGDIGETGVPGAEGPRGFPGIQGRKGEP 104

Qy         96 GECSVPPRSAFSAKRSESRVPPPSDAPLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAV 155
              || :   |||||    |: |  | : |: | ::  |:| |||  |||| | :||:||||
Db        105 GEGAYVYRSAFSV-GLETYVTIP-NMPIRFTKIFYNQQNHYDGSTGKFHCNIPGLYYFAY 162

Qy        156 HATVYRASLQFDLVKNGES-IASFFQFFGGWPKPASLSGGAMVRLEPEDQVWVQV-GVGD 213
              | |||   ::  | |  :: : :: |:       |  ||  :: ||  ||||:|| | |:
Db        163 HITVYMKDVKVSLFKKDKAMLFTYDQYQENNVDQA--SGSVLLHLEVGDQVWLQVYGEGE 220

Qy        214 YIGIYASIKTDSTFSGFLVYSD 235
                |:||    ||||:|||:| |
Db        221 RNGLYADNDNDSTFTGFLLYHD 242


RESULT 14
US-09-140-804-3
; Sequence 3, Application US/09140804
; Patent No. 6197930
; GENERAL INFORMATION:
;  APPLICANT: Sheppard, Paul O.
;  APPLICANT: Humes, Jacqueline M.
;  TITLE OF INVENTION: ADIPOCYTE-SPECIFIC PROTEIN HOMOLOGS
;  FILE REFERENCE: 97-49
;  CURRENT APPLICATION NUMBER: US/09/140,804
;  CURRENT FILING DATE: 1998-08-26
;  EARLIER APPLICATION NUMBER: 60/056,983
;  EARLIER FILING DATE: 1997-08-26
;  NUMBER OF SEQ ID NOS: 47
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 3
;   LENGTH: 244
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-140-804-3

  Query Match             32.0%;  Score 424;  DB 4;  Length 244;
  Best Local Similarity   40.5%;  Pred. No. 2.5e-31;
  Matches  106;  Conservative   26;  Mismatches   72;  Indels   58;  Gaps   10;
```

```
Qy          6 VLLLLGLAAGSPPLDDNKIPSLCPGH--------PG-------------LPGTPGHHGSQ 44
              ||||| |                 |||        ||             : | ||| |
Db          7 VLLLLAL---------------PGHDQETTTQGPGVLLPLPKGACTGWMAGIPGHPGHN 50

Qy         45 GLPGRDGRDGRDGAPGAPGEKGEGGRPGLPGPRGD---------PGPRGEAGPAGPTGPA 95
              | |||||||      | |||||| | ||| ||:||          ||||  |  |  |
Db         51 GAPGRDGRD------GTPGEKGEKGDPGLIGPKGDIGETGVPGAEGPRGFPGIQGRKGEP 104

Qy         96 GECSVPPRSAFSAKRSESRVPPPSDAPLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAV 155
              || :   |||||    |: |  | : |: | ::  |:| |||  |||| | :||:||||
Db        105 GEGAYVYRSAFSV-GLETYVTIP-NMPIRFTKIFYNQQNHYDGSTGKFHCNIPGLYYFAY 162

Qy        156 HATVYRASLQFDLVKNGES-IASFFQFFGGWPKPASLSGGAMVRLEPEDQVWVQV-GVGD 213
              | |||   ::  | |  :: : :: |:       |  ||  :: ||  ||||:|| | |:
Db        163 HITVYMKDVKVSLFKKDKAMLFTYDQYQENNVDQA--SGSVLLHLEVGDQVWLQVYGEGE 220

Qy        214 YIGIYASIKTDSTFSGFLVYSD 235
                |:||    ||||:|||:| |
Db        221 RNGLYADNDNDSTFTGFLLYHD 242


RESULT 15
US-09-336-536-20
; Sequence 20, Application US/09336536
; Patent No. 6406884
; GENERAL INFORMATION:
;  APPLICANT: Leiby, K.
;  APPLICANT: McKay, C.
;  APPLICANT: Bossone, S.
;  TITLE OF INVENTION: SECRETED PROTEINS AND USES THEREOF
;  FILE REFERENCE: 7853-144
;  CURRENT APPLICATION NUMBER: US/09/336,536
;  CURRENT FILING DATE: 1999-06-18
;  NUMBER OF SEQ ID NOS: 75
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 20
;   LENGTH: 244
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-336-536-20

  Query Match            32.0%;  Score 424;  DB 4;  Length 244;
  Best Local Similarity  40.5%;  Pred. No. 2.5e-31;
  Matches  106;  Conservative   26;  Mismatches   72;  Indels   58;  Gaps   10;

Qy          6 VLLLLGLAAGSPPLDDNKIPSLCPGH--------PG-------------LPGTPGHHGSQ 44
              ||||| |                 |||        ||             : | ||| |
Db          7 VLLLLAL---------------PGHDQETTTQGPGVLLPLPKGACTGWMAGIPGHPGHN 50

Qy         45 GLPGRDGRDGRDGAPGAPGEKGEGGRPGLPGPRGD---------PGPRGEAGPAGPTGPA 95
              | |||||||      | |||||| | ||| ||:||          ||||  |  |  |
Db         51 GAPGRDGRD------GTPGEKGEKGDPGLIGPKGDIGETGVPGAEGPRGFPGIQGRKGEP 104

Qy         96 GECSVPPRSAFSAKRSESRVPPPSDAPLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAV 155
              || :   |||||    |: |  | : |: | ::  |:| |||  |||| | :||:||||
Db        105 GEGAYVYRSAFSV-GLETYVTIP-NMPIRFTKIFYNQQNHYDGSTGKFHCNIPGLYYFAY 162

Qy        156 HATVYRASLQFDLVKNGES-IASFFQFFGGWPKPASLSGGAMVRLEPEDQVWVQV-GVGD 213
              | |||   ::  | |  :: : :: |:       |  ||  :: ||  ||||:|| | |:
Db        163 HITVYMKDVKVSLFKKDKAMLFTYDQYQENNVDQA--SGSVLLHLEVGDQVWLQVYGEGE 220

Qy        214 YIGIYASIKTDSTFSGFLVYSD 235
                |:||    ||||:|||:| |
Db        221 RNGLYADNDNDSTFTGFLLYHD 242


Search completed: June 18, 2003, 14:54:23
Job time : 30 secs
```

GenCore version 5.1.6


OM protein - protein search, using sw model

Run on: June 18, 2003, 14:54:28 ; Search time 23 Seconds
(without alignments)
438.206 Million cell updates/sec

Title: US-09-943-851A-42
Perfect score: 243
Sequence: 1 MRPLLVLLLLGLAAGSPPLD.....DSTFSGFLVYSDWHSSPVFA 243

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 112892 seqs, 41476328 residues

Word size : 0

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 243 | 100.0 | 243 | 1 | CQT5_HUMAN | Q9bxj0 homo sapien |
| 2 | 11 | 4.5 | 289 | 1 | CQT7_HUMAN | Q9bxj2 homo sapien |
| 3 | 11 | 4.5 | 369 | 1 | PSPD_BOVIN | P35246 bos taurus |
| 4 | 11 | 4.5 | 374 | 1 | PSPD_MOUSE | P50404 mus musculu |
| 5 | 11 | 4.5 | 374 | 1 | PSPD_RAT | P35248 rattus norv |
| 6 | 11 | 4.5 | 375 | 1 | PSPD_HUMAN | P35247 homo sapien |
| 7 | 11 | 4.5 | 1355 | 1 | CA21_RANCA | O42350 rana catesb |
| 8 | 11 | 4.5 | 1670 | 1 | CA34_HUMAN | Q01955 homo sapien |
| 9 | 11 | 4.5 | 2944 | 1 | CA17_HUMAN | Q02388 homo sapien |
| 10 | 10 | 4.1 | 210 | 1 | CAC2_HAECO | P16252 haemonchus |
| 11 | 10 | 4.1 | 247 | 1 | PSPA_CAVPO | P50403 cavia porce |
| 12 | 10 | 4.1 | 247 | 1 | PSPA_RABIT | P12842 oryctolagus |
| 13 | 10 | 4.1 | 248 | 1 | PSPA_CANFA | P06908 canis famil |
| 14 | 10 | 4.1 | 248 | 1 | PSPA_HUMAN | P07714 homo sapien |
| 15 | 10 | 4.1 | 248 | 1 | PSPA_MOUSE | P35242 mus musculu |
| 16 | 10 | 4.1 | 248 | 1 | PSPA_RAT | P08427 rattus norv |
| 17 | 10 | 4.1 | 249 | 1 | PSPA_PIG | P49874 sus scrofa |
| 18 | 10 | 4.1 | 526 | 1 | CA21_RABIT | Q28668 oryctolagus |
| 19 | 10 | 4.1 | 547 | 1 | CAF1_EPHMU | P18856 ephydatia m |
| 20 | 10 | 4.1 | 1049 | 1 | CA13_BOVIN | P04258 bos taurus |
| 21 | 10 | 4.1 | 1464 | 1 | CA13_MOUSE | P08121 mus musculu |
| 22 | 10 | 4.1 | 1466 | 1 | CA13_HUMAN | P02461 homo sapien |
| 23 | 9 | 3.7 | 265 | 1 | CTRC_NEIMA | P57012 neisseria m |
| 24 | 9 | 3.7 | 636 | 1 | CA13_RAT | P13941 rattus norv |
| 25 | 9 | 3.7 | 695 | 1 | APP2_MOUSE | Q06335 mus musculu |
| 26 | 9 | 3.7 | 765 | 1 | APP2_RAT | P15943 rattus norv |
| 27 | 9 | 3.7 | 1364 | 1 | CA21_BOVIN | P02465 bos taurus |
| 28 | 9 | 3.7 | 1366 | 1 | CA21_CANFA | O46392 canis famil |
| 29 | 9 | 3.7 | 1402 | 1 | IF4G_RABIT | P41110 oryctolagus |
| 30 | 9 | 3.7 | 1669 | 1 | CA14_MOUSE | P02463 mus musculu |
| 31 | 8 | 3.3 | 28 | 1 | C1QC_RAT | P31722 rattus norv |
| 32 | 8 | 3.3 | 170 | 1 | FA39_HUMAN | P49913 homo sapien |
| 33 | 8 | 3.3 | 193 | 1 | CERB_HUMAN | P23435 homo sapien |
| 34 | 8 | 3.3 | 193 | 1 | CERB_MOUSE | Q9r171 mus musculu |
| 35 | 8 | 3.3 | 201 | 1 | CERL_HUMAN | Q9ntu7 homo sapien |
| 36 | 8 | 3.3 | 224 | 1 | CERL_RAT | P98087 rattus norv |
| 37 | 8 | 3.3 | 244 | 1 | APM1_HUMAN | Q15848 homo sapien |
| 38 | 8 | 3.3 | 247 | 1 | APM1_MOUSE | Q60994 mus musculu |
| 39 | 8 | 3.3 | 301 | 1 | CC02_CAEEL | P17656 caenorhabdi |
| 40 | 8 | 3.3 | 306 | 1 | CC40_CAEEL | P34804 caenorhabdi |
| 41 | 8 | 3.3 | 316 | 1 | CC12_CAEEL | P20630 caenorhabdi |
| 42 | 8 | 3.3 | 316 | 1 | CC13_CAEEL | P20631 caenorhabdi |
| 43 | 8 | 3.3 | 329 | 1 | CQT4_HUMAN | Q9bxj3 homo sapien |
| 44 | 8 | 3.3 | 356 | 1 | VP36_CANFA | P49256 canis famil |
| 45 | 8 | 3.3 | 358 | 1 | VP36_MOUSE | Q9dbh5 mus musculu |

ALIGNMENTS

RESULT 1
CQT5_HUMAN
ID CQT5_HUMAN STANDARD; PRT; 243 AA.
AC Q9BXJ0; Q9UFX4;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Complement-c1q tumor necrosis factor-related protein 5 precursor.
GN C1QTNF5 OR CTRP5.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Sheppard P.O., Humes J.M.;
RT "Homo sapiens complement-c1q tumor necrosis factor-related protein.";
RL Submitted (DEC-2000) to the EMBL/GenBank/DDBJ databases.
RN [2]
RP SEQUENCE OF 25-243 FROM N.A.
RC TISSUE=Uterus;
RA Ottenwaelder B., Obermaier B., Mewes H.-W., Gassenhuber J.,
RA Wiemann S.;
RL Submitted (AUG-1999) to the EMBL/GenBank/DDBJ databases.
CC -!- SIMILARITY: CONTAINS 1 COLLAGENOUS DOMAIN.
CC -!- SIMILARITY: CONTAINS 1 C1Q DOMAIN.
CC ---
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC ---
DR EMBL; AF329841; AAK17965.1; -.
DR EMBL; AL110261; CAB53702.1; -.
DR Genew; HGNC:14344; C1QTNF5.
DR InterPro; IPR001073; C1q.
DR InterPro; IPR000087; Collagen.
DR Pfam; PF00386; C1q; 1.
DR Pfam; PF01391; Collagen; 1.
DR PRINTS; PR00007; COMPLEMNTC1Q.
DR SMART; SM00110; C1Q; 1.
DR PROSITE; PS01113; C1Q; FALSE_NEG.
KW Collagen; Signal.
FT SIGNAL 1 15 POTENTIAL.
FT CHAIN 16 243 COMPLEMENT-C1Q TUMOR NECROSIS FACTOR-
FT RELATED PROTEIN 5.
FT DOMAIN 30 95 COLLAGEN-LIKE.
FT DOMAIN 97 243 C1Q.
SQ SEQUENCE 243 AA; 25298 MW; 7CCDA65CDA7EB784 CRC64;

Query Match 100.0%; Score 243; DB 1; Length 243;
Best Local Similarity 100.0%; Pred. No. 1.9e-193;
Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MRPLLVLLLLGLAAGSPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MRPLLVLLLLGLAAGSPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPG 60

Qy         61 APGEKGEGGRPGLPGPRGDPGPRGEAGPAGPTGPAGECSVPPRSAFSAKRSESRVPPPSD 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 APGEKGEGGRPGLPGPRGDPGPRGEAGPAGPTGPAGECSVPPRSAFSAKRSESRVPPPSD 120

Qy        121 APLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGESIASFFQ 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 APLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGESIASFFQ 180

Qy        181 FFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWHSSP 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 FFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWHSSP 240

Qy        241 VFA 243
              |||
Db        241 VFA 243
```

```
RESULT 2
CQT7_HUMAN
ID   CQT7_HUMAN     STANDARD;      PRT;   289 AA.
AC   Q9BXJ2;
DT   15-JUN-2002 (Rel. 41, Created)
DT   15-JUN-2002 (Rel. 41, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Complement-c1q tumor necrosis factor-related protein 7 precursor.
GN   C1QTNF7 OR CTRP7.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Piddington C.S., Sheppard P.O., Bishop P., Lasser G.W.;
RT   "Homo sapiens complement-c1q tumor necrosis factor-related protein.";
RL   Submitted (DEC-2000) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Testis;
RA   Strausberg R.;
RL   Submitted (JAN-2002) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: CONTAINS 1 COLLAGENOUS DOMAIN.
CC   -!- SIMILARITY: CONTAINS 1 C1Q DOMAIN.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF329839; AAK17963.1; -.
DR   EMBL; BC022187; AAH22187.1; -.
DR   Genew; HGNC:14342; C1QTNF7.
DR   InterPro; IPR001073; C1q.
DR   InterPro; IPR000087; Collagen.
DR   Pfam; PF00386; C1q; 1.
DR   Pfam; PF01391; Collagen; 2.
DR   PRINTS; PR00007; COMPLEMNTC1Q.
DR   SMART; SM00110; C1Q; 1.
DR   PROSITE; PS01113; C1Q; 1.
KW   Collagen; Signal.
FT   SIGNAL       1     16       POTENTIAL.
FT   CHAIN       17    289       COMPLEMENT-C1Q TUMOR NECROSIS FACTOR-
FT                               RELATED PROTEIN 7.
FT   DOMAIN      38    139       COLLAGEN-LIKE.
FT   DOMAIN     141    276       C1Q.
SQ   SEQUENCE   289 AA;  30683 MW;  A61609FF86D26946 CRC64;
```

```
  Query Match             4.5%;  Score 11;  DB 1;  Length 289;
  Best Local Similarity   100.0%;  Pred. No. 0.06;
  Matches   11;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         45 GLPGRDGRDGR 55
              |||||||||||
Db         59 GLPGRDGRDGR 69
```

```
RESULT 3
PSPD_BOVIN
ID   PSPD_BOVIN     STANDARD;      PRT;   369 AA.
AC   P35246;
DT   01-FEB-1994 (Rel. 28, Created)
DT   01-FEB-1994 (Rel. 28, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   Pulmonary surfactant-associated protein D precursor (SP-D) (PSP-D).
GN   SFTPD OR SFTP4.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A., AND SEQUENCE OF 208-247.
RC   TISSUE=Lung;
RX   MEDLINE=93170856; PubMed=8436402;
RA   Lim B.L., Lu J., Reid K.B.M.;
RT   "Structural similarity between bovine conglutinin and bovine lung
RT   surfactant protein D and demonstration of liver as a site of
RT   synthesis of conglutinin.";
RL   Immunology 78:159-165(1993).
CC   -!- FUNCTION: CONTRIBUTES TO THE LUNG'S DEFENSE AGAINST INHALED
CC       MICROORGANISMS. BINDS STRONGLY MALTOSE RESIDUES AND TO A LESSER
CC       EXTENT OTHER ALPHA-GLUCOSYL MOIETIES. IT COULD PARTICIPATE IN THE
CC       EXTRACELLULAR REORGANIZATION OR TURNOVER OF PULMONARY SURFACTANT.
CC   -!- SUBUNIT: OLIGOMERIC COMPLEX OF 4 SET OF HOMOTRIMERS.
CC   -!- SUBCELLULAR LOCATION: Extracellular.
CC   -!- MISCELLANEOUS: PULMONARY SURFACTANT CONSISTS OF 90% LIPID AND 10%
CC       PROTEIN. THERE ARE 4 SURFACTANT ASSOCIATED PROTEIN: 2 COLLAGENOUS,
CC       CARBOHYDRATE-BINDING GLYCOPROTEINS (SP-A AND SP-D) AND 2 SMALL
CC       HYDROPHOBIC PROTEINS (SP-B AND SP-C).
CC   -!- SIMILARITY: CONTAINS A COLLAGENOUS DOMAIN.
CC   -!- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X75911; CAA53510.1; -.
DR   PIR; S33603; S33603.
DR   HSSP; P35247; 1B08.
DR   InterPro; IPR000087; Collagen.
DR   InterPro; IPR001304; Lectin_C.
DR   Pfam; PF00059; lectin_c; 1.
DR   Pfam; PF01391; Collagen; 2.
DR   SMART; SM00034; CLECT; 1.
DR   PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
KW   Glycoprotein; Calcium; Surface film; Gaseous exchange; Hydroxylation;
KW   Signal; Lectin; Collagen; Repeat; Coiled coil.
FT   SIGNAL       1     20       BY SIMILARITY.
FT   CHAIN       21    369       PULMONARY SURFACTANT-ASSOCIATED PROTEIN
FT                               D.
FT   DOMAIN      46    216       COLLAGEN-LIKE.
FT   DOMAIN     217    248       COILED COIL (POTENTIAL).
FT   DOMAIN     273    369       C-TYPE LECTIN (SHORT FORM).
FT   DISULFID   275    367       BY SIMILARITY.
```

```
FT   DISULFID    345    359       BY SIMILARITY.
FT   CARBOHYD     90     90       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   MOD_RES      78     78       HYDROXYLATION (BY SIMILARITY).
FT   MOD_RES      87     87       HYDROXYLATION (BY SIMILARITY).
FT   MOD_RES      96     96       HYDROXYLATION (BY SIMILARITY).
FT   MOD_RES      99     99       HYDROXYLATION (BY SIMILARITY).
FT   MOD_RES     165    165       HYDROXYLATION (BY SIMILARITY).
FT   MOD_RES     171    171       HYDROXYLATION (BY SIMILARITY).
SQ   SEQUENCE   369 AA;  37361 MW;  07D88B24E0AEB2E3 CRC64;

  Query Match             4.5%;  Score 11;  DB 1;  Length 369;
  Best Local Similarity   100.0%;  Pred. No. 0.073;
  Matches   11;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         45 GLPGRDGRDGR 55
              ||||||||||| 
Db         46 GLPGRDGRDGR 56


RESULT 4
PSPD_MOUSE
ID   PSPD_MOUSE     STANDARD;      PRT;   374 AA.
AC   P50404;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Pulmonary surfactant-associated protein D precursor (SP-D) (PSP-D).
GN   SFTPD OR SFTP4.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6 X CBA; TISSUE=Lung;
RX   MEDLINE=96094460; PubMed=7499852;
RA   Motwani M., White R.A., Guo N., Dowler L.L., Tauber A.I., Sastry K.N.;
RT   "Mouse surfactant protein-D. cDNA cloning, characterization, and gene
RT   localization to chromosome 14.";
RL   J. Immunol. 155:5671-5677(1995).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=129/Sv;
RX   MEDLINE=99244602; PubMed=10226065;
RA   Lawson P.R., Perkins V.C., Holmskov U., Reid K.B.;
RT   "Genomic organization of the mouse gene for lung surfactant protein
RT   D.";
RL   Am. J. Respir. Cell Mol. Biol. 20:953-963(1999).
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Fisher J.H., Sheftelyevich V.V.;
RT   "Surfactant protein-D regulates surfactant phospholipid homeostasis in
RT   vivo.";
RL   Submitted (OCT-1999) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE FROM N.A.
RA   Strausberg R.;
RL   Submitted (FEB-2001) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: CONTRIBUTES TO THE LUNG'S DEFENSE AGAINST INHALED
CC       MICROORGANISMS. BINDS STRONGLY MALTOSE RESIDUES AND TO A LESSER
CC       EXTENT OTHER ALPHA-GLUCOSYL MOIETIES. IT COULD PARTICIPATE IN THE
CC       EXTRACELLULAR REORGANIZATION OR TURNOVER OF PULMONARY SURFACTANT.
CC   -!- SUBUNIT: OLIGOMERIC COMPLEX OF 4 SET OF HOMOTRIMERS.
CC   -!- SUBCELLULAR LOCATION: Extracellular.
CC   -!- MISCELLANEOUS: PULMONARY SURFACTANT CONSISTS OF 90% LIPID AND 10%
CC       PROTEIN. THERE ARE 4 SURFACTANT ASSOCIATED PROTEIN: 2 COLLAGENOUS,
CC       CARBOHYDRATE-BINDING GLYCOPROTEINS (SP-A AND SP-D) AND 2 SMALL
CC       HYDROPHOBIC PROTEINS (SP-B AND SP-C).
CC   -!- SIMILARITY: CONTAINS A COLLAGENOUS DOMAIN.
CC   -!- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L40156; AAA92021.1; -.
DR   EMBL; AF047742; AAD31380.1; -.
DR   EMBL; AF047741; AAD31380.1; JOINED.
DR   EMBL; AF192134; AAF15277.1; -.
DR   EMBL; BC003705; AAH03705.1; -.
DR   HSSP; P35247; 1B08.
DR   MGD; MGI:109515; Sftpd.
DR   InterPro; IPR000087; Collagen.
DR   InterPro; IPR001304; Lectin_C.
DR   Pfam; PF00059; lectin_c; 1.
DR   Pfam; PF01391; Collagen; 3.
DR   SMART; SM00034; CLECT; 1.
DR   PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
KW   Glycoprotein; Calcium; Surface film; Gaseous exchange; Hydroxylation;
KW   Signal; Lectin; Collagen; Repeat; Coiled coil.
FT   SIGNAL        1     19       BY SIMILARITY.
FT   CHAIN        20    374       PULMONARY SURFACTANT-ASSOCIATED PROTEIN
FT                                D.
FT   DOMAIN       45    221       COLLAGEN-LIKE.
FT   DOMAIN      222    253       COILED COIL (POTENTIAL).
FT   DOMAIN      278    374       C-TYPE LECTIN (SHORT FORM).
FT   DISULFID    280    372       BY SIMILARITY.
FT   DISULFID    350    364       BY SIMILARITY.
FT   CARBOHYD     89     89       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   374 AA;  37688 MW;  FE034261263F43E4 CRC64;

  Query Match             4.5%;  Score 11;  DB 1;  Length 374;
  Best Local Similarity   100.0%;  Pred. No. 0.074;
  Matches   11;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         45 GLPGRDGRDGR 55
              ||||||||||| 
Db         45 GLPGRDGRDGR 55


RESULT 5
PSPD_RAT
ID   PSPD_RAT       STANDARD;      PRT;   374 AA.
AC   P35248;
DT   01-FEB-1994 (Rel. 28, Created)
DT   01-FEB-1994 (Rel. 28, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   Pulmonary surfactant-associated protein D precursor (SP-D) (PSP-D)
DE   (CP4).
GN   SFTPD OR SFTP4.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A., AND SEQUENCE OF 20-33.
RC   TISSUE=Lung;
RX   MEDLINE=92112913; PubMed=1370483;
RA   Shimizu H., Fisher J.H., Papst P., Benson B., Lau K., Mason R.J.,
RA   Voelker D.R.;
RT   "Primary structure of rat pulmonary surfactant protein D. cDNA and
RT   deduced amino acid sequence.";
RL   J. Biol. Chem. 267:1853-1857(1992).
RN   [2]
RP   SEQUENCE OF 73-95 AND 153-180.
RC   TISSUE=Lung;
RX   MEDLINE=90001186; PubMed=2675969;
RA   Persson A., Chang D., Rust K., Moxley M., Longmore W., Crouch E.;
RT   "Purification and biochemical characterization of CP4 (SP-D), a
RT   collagenous surfactant-associated protein.";
RL   Biochemistry 28:6361-6367(1989).
```

```
CC   -!- FUNCTION: CONTRIBUTES TO THE LUNG'S DEFENSE AGAINST INHALED
CC       MICROORGANISMS. BINDS STRONGLY MALTOSE RESIDUES AND TO A LESSER
CC       EXTENT OTHER ALPHA-GLUCOSYL MOIETIES. IT COULD PARTICIPATE IN THE
CC       EXTRACELLULAR REORGANIZATION OR TURNOVER OF PULMONARY SURFACTANT.
CC   -!- SUBUNIT: OLIGOMERIC COMPLEX OF 4 SET OF HOMOTRIMERS.
CC   -!- SUBCELLULAR LOCATION: Extracellular.
CC   -!- MISCELLANEOUS: PULMONARY SURFACTANT CONSISTS OF 90% LIPID AND 10%
CC       PROTEIN. THERE ARE 4 SURFACTANT ASSOCIATED PROTEIN: 2 COLLAGENOUS,
CC       CARBOHYDRATE-BINDING GLYCOPROTEINS (SP-A AND SP-D) AND 2 SMALL
CC       HYDROPHOBIC PROTEINS (SP-B AND SP-C).
CC   -!- SIMILARITY: CONTAINS A COLLAGENOUS DOMAIN.
CC   -!- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; M81231; AAA42170.1; -.
DR   PIR; A42046; A42046.
DR   HSSP; P35247; 1B08.
DR   InterPro; IPR000087; Collagen.
DR   InterPro; IPR001304; Lectin_C.
DR   Pfam; PF00059; lectin_c; 1.
DR   Pfam; PF01391; Collagen; 3.
DR   SMART; SM00034; CLECT; 1.
DR   PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
KW   Glycoprotein; Calcium; Surface film; Gaseous exchange; Hydroxylation;
KW   Signal; Lectin; Collagen; Repeat; Coiled coil.
FT   SIGNAL        1     19
FT   CHAIN        20    374       PULMONARY SURFACTANT-ASSOCIATED PROTEIN
FT                                 D.
FT   DOMAIN       45    221       COLLAGEN-LIKE.
FT   DOMAIN      222    253       COILED COIL (POTENTIAL).
FT   DOMAIN      278    374       C-TYPE LECTIN (SHORT FORM).
FT   DISULFID    280    372       BY SIMILARITY.
FT   DISULFID    350    364       BY SIMILARITY.
FT   CARBOHYD     89     89       N-LINKED (GLCNAC. . .).
FT   MOD_RES      77     77       HYDROXYLATION.
FT   MOD_RES      86     86       HYDROXYLATION.
FT   MOD_RES      95     95       HYDROXYLATION.
FT   MOD_RES      98     98       HYDROXYLATION.
FT   MOD_RES     170    170       HYDROXYLATION.
FT   MOD_RES     176    176       HYDROXYLATION.
FT   CONFLICT     89     89       N -> E (IN REF. 2).
FT   CONFLICT    164    164       K -> C (IN REF. 2).
SQ   SEQUENCE   374 AA;  37561 MW;  DB2BB5E399DB4A3C CRC64;

  Query Match             4.5%;  Score 11;  DB 1;  Length 374;
  Best Local Similarity   100.0%;  Pred. No. 0.074;
  Matches   11;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         45 GLPGRDGRDGR 55
              ||||||||||
Db         45 GLPGRDGRDGR 55


RESULT 6
PSPD_HUMAN
ID   PSPD_HUMAN     STANDARD;      PRT;   375 AA.
AC   P35247;
DT   01-FEB-1994 (Rel. 28, Created)
DT   01-FEB-1994 (Rel. 28, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Pulmonary surfactant-associated protein D precursor (SP-D) (PSP-D).
GN   SFTPD OR SFTP4 OR PSPD.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Placenta;
RX   MEDLINE=93155122; PubMed=8428971;
RA   Crouch E., Rust K., Veile R., Donis-Keller H., Grosso L.;
RT   "Genomic organization of human surfactant protein D (SP-D). SP-D is
RT   encoded on chromosome 10q22.2-23.1.";
RL   J. Biol. Chem. 268:2976-2983(1993).
RN   [2]
RP   SEQUENCE FROM N.A., AND SEQUENCE OF 214-243.
RC   TISSUE=Lung, and Amniotic fluid;
RX   MEDLINE=92322003; PubMed=1339284;
RA   Lu J., Willis A.C., Reid K.B.M.;
RT   "Purification, characterization and cDNA cloning of human lung
RT   surfactant protein D.";
RL   Biochem. J. 284:795-802(1992).
RN   [3]
RP   SEQUENCE OF 60-375 FROM N.A., AND PARTIAL SEQUENCE.
RC   TISSUE=Lung;
RX   MEDLINE=91378578; PubMed=1898081;
RA   Rust K., Grosso L., Zhang V., Chang D., Persson A., Longmore W.,
RA   Cai G.-Z., Crouch E.;
RT   "Human surfactant protein D: SP-D contains a C-type lectin
RT   carbohydrate recognition domain.";
RL   Arch. Biochem. Biophys. 290:116-126(1991).
RN   [4]
RP   X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS).
RX   MEDLINE=99197291; PubMed=10368295;
RA   Hakansson K., Lim N.K., Hoppe H.-J., Reid K.B.M.;
RT   "Crystal structure of the trimeric alpha-helical coiled-coil and the
RT   three lectin domains of human lung surfactant protein D.";
RL   Structure 7:255-264(1999).
CC   -!- FUNCTION: CONTRIBUTES TO THE LUNG'S DEFENSE AGAINST INHALED
CC       MICROORGANISMS. BINDS STRONGLY MALTOSE RESIDUES AND TO A LESSER
CC       EXTENT OTHER ALPHA-GLUCOSYL MOIETIES. IT COULD PARTICIPATE IN THE
CC       EXTRACELLULAR REORGANIZATION OR TURNOVER OF PULMONARY SURFACTANT.
CC   -!- SUBUNIT: OLIGOMERIC COMPLEX OF 4 SET OF HOMOTRIMERS.
CC   -!- SUBCELLULAR LOCATION: Extracellular.
CC   -!- PTM: The N-terminus is blocked.
CC   -!- MISCELLANEOUS: PULMONARY SURFACTANT CONSISTS OF 90% LIPID AND 10%
CC       PROTEIN. THERE ARE 4 SURFACTANT ASSOCIATED PROTEIN: 2 COLLAGENOUS,
CC       CARBOHYDRATE-BINDING GLYCOPROTEINS (SP-A AND SP-D) AND 2 SMALL
CC       HYDROPHOBIC PROTEINS (SP-B AND SP-C).
CC   -!- SIMILARITY: CONTAINS A COLLAGENOUS DOMAIN.
CC   -!- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; L05485; AAB59450.1; -.
DR   EMBL; L05483; AAB59450.1; JOINED.
DR   EMBL; L05484; AAB59450.1; JOINED.
DR   EMBL; X65018; CAA46152.1; -.
DR   PIR; A45225; A45225.
DR   PIR; S18382; S18382.
DR   PDB; 1B08; 29-NOV-99.
DR   Genew; HGNC:10803; SFTPD.
DR   MIM; 178635; -.
DR   InterPro; IPR000087; Collagen.
DR   InterPro; IPR001304; Lectin_C.
DR   Pfam; PF00059; lectin_c; 1.
DR   Pfam; PF01391; Collagen; 4.
DR   SMART; SM00034; CLECT; 1.
DR   PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
KW   Glycoprotein; Calcium; Surface film; Gaseous exchange; Hydroxylation;
KW   Signal; Lectin; Collagen; Repeat; Coiled coil; 3D-structure.
```

```
FT   SIGNAL        1     20       BY SIMILARITY.
FT   CHAIN        21    375       PULMONARY SURFACTANT-ASSOCIATED PROTEIN
FT                                 D.
FT   DOMAIN       46    222       COLLAGEN-LIKE.
FT   DOMAIN      223    252       COILED COIL (POTENTIAL).
FT   DOMAIN      279    375       C-TYPE LECTIN (SHORT FORM).
FT   DISULFID    281    373
FT   DISULFID    351    365
FT   CARBOHYD     90     90       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   MOD_RES      78     78       HYDROXYLATION (BY SIMILARITY).
FT   MOD_RES      87     87       HYDROXYLATION (BY SIMILARITY).
FT   MOD_RES      96     96       HYDROXYLATION (BY SIMILARITY).
FT   MOD_RES      99     99       HYDROXYLATION (BY SIMILARITY).
FT   MOD_RES     171    171       HYDROXYLATION (BY SIMILARITY).
FT   MOD_RES     177    177       HYDROXYLATION (BY SIMILARITY).
FT   CONFLICT     31     31       M -> T (IN REF. 2).
FT   CONFLICT     59     59       P -> F (IN REF. 3).
FT   CONFLICT    122    122       A -> P (IN REF. 2).
FT   CONFLICT    180    180       T -> A (IN REF. 2).
FT   CONFLICT    206    206       D -> P (IN REF. 3).
FT   CONFLICT    374    374       E -> EH (IN REF. 3).
SQ   SEQUENCE   375 AA;  37702 MW;  2986B2699FC01A6A CRC64;

  Query Match             4.5%;  Score 11;  DB 1;  Length 375;
  Best Local Similarity   100.0%;  Pred. No. 0.074;
  Matches   11;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         45 GLPGRDGRDGR 55
              ||||||||||| 
Db         46 GLPGRDGRDGR 56


RESULT 7
CA21_RANCA
ID   CA21_RANCA     STANDARD;      PRT;  1355 AA.
AC   O42350;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Collagen alpha 2(I) chain precursor.
GN   COL1A2.
OS   Rana catesbeiana (Bull frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Neobatrachia; Ranoidea; Ranidae; Rana.
OX   NCBI_TaxID=8400;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Tail;
RX   MEDLINE=97417499; PubMed=9272872;
RA   Asahina K., Oofusa K., Obara M., Yoshizato K.;
RT   "Cloning and characterization of the full length cDNA encoding alpha2
RT   type I collagen of bullfrog Rana catesbeiana.";
RL   Gene 194:283-289(1997).
CC   -!- FUNCTION: TYPE I COLLAGEN IS A MEMBER OF GROUP I COLLAGEN
CC       (FIBRILLAR FORMING COLLAGEN).
CC   -!- SUBUNIT: TRIMERS OF ONE ALPHA 2(I) AND TWO ALPHA 1(I) CHAINS.
CC   -!- TISSUE SPECIFICITY: FORMS THE FIBRILS OF TENDON, LIGAMENTS AND
CC       BONES. IN BONES THE FIBRILS ARE MINERALIZED WITH CALCIUM
CC       HYDROXYAPATITE.
CC   -!- PTM: PROLINES AT THE THIRD POSITION OF THE TRIPEPTIDE REPEATING
CC       UNIT (G-X-Y) ARE HYDROXYLATED IN SOME OR ALL OF THE CHAINS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; D88764; BAA22380.1; -.
DR   InterPro; IPR000087; Collagen.
DR   InterPro; IPR000885; Fib_collagen_C.
DR   Pfam; PF01391; Collagen; 18.
DR   Pfam; PF01410; COLFI; 1.
DR   ProDom; PD000007; Collagen; 3.
DR   ProDom; PD002078; Fib_collagen_C; 1.
DR   SMART; SM00038; COLFI; 1.
KW   Extracellular matrix; Connective tissue; Repeat; Hydroxylation;
KW   Glycoprotein; Collagen; Signal.
FT   SIGNAL        1     22       POTENTIAL.
FT   PROPEP       23     82       AMINO-TERMINAL PROPEPTIDE (POTENTIAL).
FT   CHAIN        83   1093       COLLAGEN ALPHA 2(I) CHAIN.
FT   PROPEP     1094   1355       CARBOXYL-TERMINAL PROPEPTIDE
FT                                (BY SIMILARITY).
FT   CARBOHYD   1206   1206       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1256   1256       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   1355 AA;  127643 MW;  CB793AD5D6F41D2A CRC64;

  Query Match             4.5%;  Score 11;  DB 1;  Length 1355;
  Best Local Similarity   100.0%;  Pred. No. 0.21;
  Matches   11;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         83 RGEAGPAGPTG 93
              ||||||||||| 
Db        705 RGEAGPAGPTG 715


RESULT 8
CA34_HUMAN
ID   CA34_HUMAN     STANDARD;      PRT;  1670 AA.
AC   Q01955; Q9BQT2;
DT   01-OCT-1996 (Rel. 34, Created)
DT   15-JUL-1999 (Rel. 38, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Collagen alpha 3(IV) chain precursor (Goodpasture antigen).
GN   COL4A3.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Kidney;
RX   MEDLINE=94364994; PubMed=8083201;
RA   Mariyama M., Leinonen A., Mochizuki T., Tryggvason K., Reeders S.T.;
RT   "Complete primary structure of the human alpha 3(IV) collagen chain.
RT   Coexpression of the alpha 3(IV) and alpha 4(IV) collagen chains in
RT   human tissues.";
RL   J. Biol. Chem. 269:23013-23017(1994).
RN   [2]
RP   REVISIONS.
RA   Leinonen A.;
RL   Submitted (OCT-1998) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A., VARIANTS AS E-297; R-407; R-640; R-1167; E-1207;
RP   Q-1215; S-1277; T-1330; E-1334; E-1347 AND C-1661, AND VARIANTS R-43;
RP   E-162; Y-326; H-408; R-451; L-574; E-1269 AND P-1474.
RX   MEDLINE=21064696; PubMed=11134255;
RA   Heidet L., Arrondel C., Forestier L., Cohen-Solal L., Mollet G.,
RA   Gutierrez B., Stavrou C., Gubler M.C., Antignac C.;
RT   "Structure of the human type IV collagen gene COL4A3 and mutations in
RT   autosomal Alport syndrome.";
RL   J. Am. Soc. Nephrol. 12:97-106(2001).
RN   [4]
RP   SEQUENCE OF 1386-1670 FROM N.A., AND PARTIAL SEQUENCE.
RX   MEDLINE=93015826; PubMed=1400291;
RA   Quinones S., Bernal D., Garcia-Sogo M., Elena S.F., Saus J.;
RT   "Exon/intron structure of the human alpha 3(IV) gene encompassing the
RT   Goodpasture antigen (alpha 3(IV)NC1). Identification of a potentially
RT   antigenic region at the triple helix/NC1 domain junction.";
RL   J. Biol. Chem. 267:19780-19784(1992).
RN   [5]
RP   SEQUENCE OF 1453-1670 FROM N.A.
RX   MEDLINE=91353570; PubMed=1882840;
RA   Morrison K.E., Mariyama M., Yang-Feng T.L., Reeders S.T.;
```

```
RT   "Sequence and localization of a partial cDNA encoding the human alpha
RT   3 chain of type IV collagen.";
RL   Am. J. Hum. Genet. 49:545-554(1991).
RN   [6]
RP   SEQUENCE OF 1331-1670 FROM N.A.
RC   TISSUE=Kidney;
RX   MEDLINE=92147878; PubMed=1737849;
RA   Turner N., Mason P.J., Brown R., Fox M., Povey S., Rees A.,
RA   Pusey C.D.;
RT   "Molecular cloning of the human Goodpasture antigen demonstrates it
RT   to be the alpha 3 chain of type IV collagen.";
RL   J. Clin. Invest. 89:592-601(1992).
RN   [7]
RP   SEQUENCE OF 1644-1670 FROM N.A.
RC   TISSUE=Kidney;
RA   Ding J.;
RL   Submitted (JAN-1993) to the EMBL/GenBank/DDBJ databases.
RN   [8]
RP   SEQUENCE OF 1439-1670, AND ALTERNATIVE SPLICING.
RC   TISSUE=Kidney;
RX   MEDLINE=94124597; PubMed=8294492;
RA   Feng L., Xia Y., Wilson C.B.;
RT   "Alternative splicing of the NC1 domain of the human alpha 3(IV)
RT   collagen gene. Differential expression of mRNA transcripts that
RT   predict three protein variants with distinct carboxyl regions.";
RL   J. Biol. Chem. 269:2342-2348(1994).
RN   [9]
RP   SEQUENCE OF 1-29 FROM N.A.
RX   MEDLINE=98196854; PubMed=9537506;
RA   Momota R., Sugimoto M., Oohashi T., Kigasawa K., Yoshioka H.,
RA   Ninomiya Y.;
RT   "Two genes, COL4A3 and COL4A4 coding for the human alpha3(IV) and
RT   alpha4(IV) collagen chains are arranged head-to-head on chromosome
RT   2q36.";
RL   FEBS Lett. 424:11-16(1998).
RN   [10]
RP   ALTERNATIVE SPLICING.
RX   MEDLINE=93280184; PubMed=8505332;
RA   Bernal D., Quinones S., Saus J.;
RT   "The human mRNA encoding the Goodpasture antigen is alternatively
RT   spliced.";
RL   J. Biol. Chem. 268:12090-12094(1993).
RN   [11]
RP   VARIANT PRO-1474.
RX   MEDLINE=95078827; PubMed=7987301;
RA   Lemmink H.H., Mochizuki T., van den Heuvel L.P.W.J., Schroeder C.H.,
RA   Barrientos A., Monnens L.A.H., van Oost B.A., Brunner H.G.,
RA   Reeders S.T., Smeets H.J.M.;
RT   "Mutations in the type IV collagen alpha 3 (COL4A3) gene in autosomal
RT   recessive Alport syndrome.";
RL   Hum. Mol. Genet. 3:1269-1273(1994).
CC   -!- FUNCTION: TYPE IV COLLAGEN IS THE MAJOR STRUCTURAL COMPONENT OF
CC       GLOMERULAR BASEMENT MEMBRANES (GBM), FORMING A 'CHICKEN-WIRE'
CC       MESHWORK TOGETHER WITH LAMININS, PROTEOGLYCANS AND ENTACTIN/
CC       NIDOGEN.
CC   -!- SUBUNIT: THERE ARE SIX TYPE IV COLLAGEN ISOFORMS, ALPHA 1(IV)-
CC       ALPHA 6(IV), EACH OF WHICH CAN FORM A TRIPLE HELIX STRUCTURE
CC       WITH 2 OTHER CHAINS TO GENERATE TYPE IV COLLAGEN NETWORK.
CC   -!- SUBCELLULAR LOCATION: CELL SURFACE (POTENTIAL).
CC   -!- ALTERNATIVE PRODUCTS: AT LEAST 3 ISOFORMS; 1 (SHOWN HERE), 2/V AND
CC       3/L5; ARE PRODUCED BY ALTERNATIVE SPLICING. THEY DIFFER IN THEIR
CC       C-TERMINAL NC1 DOMAINS.
CC   -!- TISSUE SPECIFICITY: ALPHA 3 AND ALPHA 4 TYPE IV COLLAGENS ARE
CC       COLOCALIZED AND PRESENT ONLY IN BASEMENT MEMBRANES OF KIDNEY, EYE,
CC       COCHLEA, LUNG AND BRAIN.
CC   -!- DOMAIN: ALPHA CHAINS OF TYPE IV COLLAGEN HAVE A NONCOLLAGENOUS
CC       DOMAIN (NC1) AT THEIR C-TERMINUS, FREQUENT INTERRUPTIONS OF THE
CC       G-X-Y REPEATS IN THE LONG CENTRAL TRIPLE-HELICAL DOMAIN (WHICH MAY
CC       CAUSE FLEXIBILITY IN THE TRIPLE HELIX), AND A SHORT N-TERMINAL
CC       TRIPLE-HELICAL 7S DOMAIN.
CC   -!- PTM: PROLINES AT THE THIRD POSITION OF THE TRIPEPTIDE REPEATING
CC       UNIT (G-X-Y) ARE HYDROXYLATED IN SOME OR ALL OF THE CHAINS.
CC   -!- PTM: THE ALTERNATIVE SPLICED FORM V CONTAINS AN ADDITIONAL
CC       N-LINKED GLYCOSYLATION SITE.
CC   -!- PTM: TYPE IV COLLAGENS CONTAIN NUMEROUS CYSTEINE RESIDUES WHICH
CC       ARE INVOLVED IN INTER- AND INTRAMOLECULAR DISULFIDE BONDING. 12 OF
CC       THESE, LOCATED IN THE NC1 DOMAIN, ARE CONSERVED IN ALL KNOWN TYPE
CC       IV COLLAGENS.
CC   -!- PTM: Phosphorylated by the Goodpasture antigen-binding protein.
CC   -!- DISEASE: ANTIBODIES AGAINST THE NC1 DOMAIN OF ALPHA3(IV) MEDIATE
CC       THE HUMAN AUTOIMMUNE DISEASE, GOODPASTURE SYNDROME, WHICH IS
CC       CHARACTERIZED BY HEMATURIA AND PULMONARY HEMORRHAGE.
CC   -!- DISEASE: DEFECTS IN COL4A3 ARE ASSOCIATED WITH THE TYPE I
CC       AUTOSOMAL RECESSIVE FORM OF ALPORT SYNDROME, AN HEREDITARY
CC       GLOMERULONEPHROPATHY CHARACTERIZED BY PROGRESSIVE RENAL FAILURE,
CC       HEMATURIA AND DEAFNESS. THE RECESSIVE FORM OCCURS EQUALLY BETWEEN
CC       MALES AND FEMALES.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X80031; CAA56335.1; -.
DR   EMBL; AJ288487; CAC36101.1; -.
DR   EMBL; AJ288488; CAC36101.1; JOINED.
DR   EMBL; AJ288489; CAC36101.1; JOINED.
DR   EMBL; AJ288490; CAC36101.1; JOINED.
DR   EMBL; AJ288491; CAC36101.1; JOINED.
DR   EMBL; AJ288492; CAC36101.1; JOINED.
DR   EMBL; AJ288493; CAC36101.1; JOINED.
DR   EMBL; AJ288494; CAC36101.1; JOINED.
DR   EMBL; AJ288495; CAC36101.1; JOINED.
DR   EMBL; AJ288496; CAC36101.1; JOINED.
DR   EMBL; AJ288497; CAC36101.1; JOINED.
DR   EMBL; AJ288498; CAC36101.1; JOINED.
DR   EMBL; AJ288499; CAC36101.1; JOINED.
DR   EMBL; AJ288500; CAC36101.1; JOINED.
DR   EMBL; AJ288501; CAC36101.1; JOINED.
DR   EMBL; AJ288502; CAC36101.1; JOINED.
DR   EMBL; AJ288503; CAC36101.1; JOINED.
DR   EMBL; AJ288504; CAC36101.1; JOINED.
DR   EMBL; AJ288505; CAC36101.1; JOINED.
DR   EMBL; AJ288506; CAC36101.1; JOINED.
DR   EMBL; AJ288507; CAC36101.1; JOINED.
DR   EMBL; AJ288508; CAC36101.1; JOINED.
DR   EMBL; AJ288509; CAC36101.1; JOINED.
DR   EMBL; AJ288510; CAC36101.1; JOINED.
DR   EMBL; AJ288511; CAC36101.1; JOINED.
DR   EMBL; AJ288512; CAC36101.1; JOINED.
DR   EMBL; AJ288513; CAC36101.1; JOINED.
DR   EMBL; AJ288514; CAC36101.1; JOINED.
DR   EMBL; AJ288515; CAC36101.1; JOINED.
DR   EMBL; AJ288516; CAC36101.1; JOINED.
DR   EMBL; AJ288517; CAC36101.1; JOINED.
DR   EMBL; AJ288518; CAC36101.1; JOINED.
DR   EMBL; AJ288519; CAC36101.1; JOINED.
DR   EMBL; AJ288520; CAC36101.1; JOINED.
DR   EMBL; AJ288521; CAC36101.1; JOINED.
DR   EMBL; AJ288522; CAC36101.1; JOINED.
DR   EMBL; AJ288523; CAC36101.1; JOINED.
DR   EMBL; AJ288524; CAC36101.1; JOINED.
DR   EMBL; AJ288525; CAC36101.1; JOINED.
DR   EMBL; AJ288526; CAC36101.1; JOINED.
DR   EMBL; AJ288527; CAC36101.1; JOINED.
DR   EMBL; AJ288528; CAC36101.1; JOINED.
DR   EMBL; AJ288529; CAC36101.1; JOINED.
DR   EMBL; AJ288530; CAC36101.1; JOINED.
DR   EMBL; AJ288531; CAC36101.1; JOINED.
DR   EMBL; AJ288532; CAC36101.1; JOINED.
DR   EMBL; AJ288533; CAC36101.1; JOINED.
DR   EMBL; AJ288534; CAC36101.1; JOINED.
DR   EMBL; AJ288535; CAC36101.1; JOINED.
```

```
DR   EMBL; AJ288536; CAC36101.1; JOINED.
DR   EMBL; AJ288537; CAC36101.1; JOINED.
DR   EMBL; AJ288538; CAC36101.1; JOINED.
DR   EMBL; M92993; AAA21610.1; -.
DR   EMBL; S55790; AAB19637.1; -.
DR   EMBL; M81379; AAA51556.1; -.
DR   EMBL; L08650; AAA52044.1; -.
DR   EMBL; U02519; AAA18942.1; -.

  Query Match              4.5%;  Score 11;  DB 1;  Length 1670;
  Best Local Similarity   100.0%;  Pred. No. 0.24;
  Matches   11;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         71 PGLPGPRGDPG 81
              ||||||||||| 
Db       1300 PGLPGPRGDPG 1310


RESULT 9
CA17_HUMAN
ID   CA17_HUMAN     STANDARD;      PRT;  2944 AA.
AC   Q02388; Q14054; Q16507;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-FEB-1996 (Rel. 33, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Collagen alpha 1(VII) chain precursor (Long-chain collagen) (LC
DE   collagen).
GN   COL7A1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94327588; PubMed=8051117;
RA   Christiano A.M., Greenspan D.S., Lee S., Uitto J.;
RT   "Cloning of human type VII collagen. Complete primary sequence of the
RT   alpha 1(VII) chain and identification of intragenic polymorphisms.";
RL   J. Biol. Chem. 269:20256-20262(1994).
RN   [2]
RP   SEQUENCE OF 128-1493 FROM N.A., AND PARTIAL SEQUENCE.
RX   MEDLINE=93338437; PubMed=1307247;
RA   Christiano A.M., Rosenbaum L.M., Chung-Honet L.C., Parente M.G.,
RA   Woodley D.T., Pan T.C., Zhang R.Z., Chu M.-L., Burgeson R.E.,
RA   Uitto J.;
RT   "The large non-collagenous domain (NC-1) of type VII collagen is
RT   amino-terminal and chimeric. Homology to cartilage matrix protein,
RT   the type III domains of fibronectin and the A domains of von
RT   Willebrand factor.";
RL   Hum. Mol. Genet. 1:475-481(1992).
RN   [3]
RP   SEQUENCE OF 815-1439 FROM N.A.
RX   MEDLINE=91334380; PubMed=1871109;
RA   Parente M.G., Chung L.C., Ryynaenen J., Woodley D.T., Wynn K.W.,
RA   Bauer E.A., Mattei M.-G., Chu M.-L., Uitto J.;
RT   "Human type VII collagen: cDNA cloning and chromosomal mapping of the
RT   gene.";
RL   Proc. Natl. Acad. Sci. U.S.A. 88:6931-6935(1991).
RN   [4]
RP   SEQUENCE OF 369-1255 FROM N.A.
RX   MEDLINE=93107742; PubMed=1469284;
RA   Gammon W.R., Abernethy M.L., Padilla K.M., Prisayanh P.S.,
RA   Cook M.E., Wright J., Briggaman R.A., Hunt S.W. III;
RT   "Noncollagenous (NC1) domain of collagen VII resembles multidomain
RT   adhesion proteins involved in tissue-specific organization of
RT   extracellular matrix.";
RL   J. Invest. Dermatol. 99:691-696(1992).
RN   [5]
RP   SEQUENCE OF 340-675 FROM N.A.
RC   TISSUE=Keratinocytes;
RX   MEDLINE=92231902; PubMed=1567409;
RA   Tanaka T., Takahashi K., Furukawa F., Imamura S.;
RT   "Molecular cloning and characterization of type VII collagen cDNA.";
RL   Biochem. Biophys. Res. Commun. 183:958-963(1992).
RN   [6]
RP   SEQUENCE OF 2395-2944 FROM N.A.
RX   MEDLINE=93271985; PubMed=8499916;
RA   Greenspan D.S.;
RT   "The carboxyl-terminal half of type VII collagen, including the non-
RT   collagenous NC-2 domain and intron/exon organization of the
RT   corresponding region of the COL7A1 gene.";
RL   Hum. Mol. Genet. 2:273-278(1993).
RN   [7]
RP   SEQUENCE OF 1-87 FROM N.A.
RC   TISSUE=Placenta;
RX   MEDLINE=94375010; PubMed=8088784;
RA   Christiano A.M., Hoffman G.G., Chung-Honet L.C., Lee S., Cheng W.,
RA   Uitto J., Greenspan D.S.;
RT   "Structural organization of the human type VII collagen gene (COL7A1),
RT   composed of more exons than any previously characterized gene.";
RL   Genomics 21:169-179(1994).
RN   [8]
RP   REVIEW ON DEB VARIANTS.
RX   MEDLINE=98041696; PubMed=9375848;
RA   Jaervikallio A., Pulkkinen L., Uitto J.;
RT   "Molecular basis of dystrophic epidermolysis bullosa: mutations in
RT   the type VII collagen gene (COL7A1).";
RL   Hum. Mutat. 10:338-347(1997).
RN   [9]
RP   VARIANT RDEB LYS-2798.
RX   MEDLINE=93291877; PubMed=8513326;
RA   Christiano A.M., Greenspan D.S., Hoffman G.G., Zhang X., Tamai Y.,
RA   Lin A.N., Dietz H.C., Hovnanian A., Uitto J.;
RT   "A missense mutation in type VII collagen in two affected siblings
RT   with recessive dystrophic epidermolysis bullosa.";
RL   Nat. Genet. 4:62-66(1993).
RN   [10]
RP   VARIANT DDEB SER-2040.
RX   MEDLINE=94224777; PubMed=8170945;
RA   Christiano A.M., Ryynaenen M., Uitto J.;
RT   "Dominant dystrophic epidermolysis bullosa: identification of a
RT   Gly-->Ser substitution in the triple-helical domain of type VII
RT   collagen.";
RL   Proc. Natl. Acad. Sci. U.S.A. 91:3549-3553(1994).
RN   [11]
RP   VARIANT PEB-DDEB CYS-2623.
RX   MEDLINE=96081220; PubMed=8541842;
RA   Christiano A.M., Lee J.Y.-Y., Chen W.J., Laforgia S., Uitto J.;
RT   "Pretibial epidermolysis bullosa: genetic linkage to COL7A1 and
RT   identification of a glycine-to-cysteine substitution in the triple-
RT   helical domain of type VII collagen.";
RL   Hum. Mol. Genet. 4:1579-1583(1995).
RN   [12]
RP   VARIANT DDEB ARG-2043.
RX   MEDLINE=95164985; PubMed=7861014;
RA   Christiano A.M., Morricone A., Paradisi M., Angelo C., Mazzanti C.,
RA   Cavalieri R., Uitto J.;
RT   "A glycine-to-arginine substitution in the triple-helical domain of
RT   type VII collagen in a family with dominant dystrophic epidermolysis
RT   bullosa.";
RL   J. Invest. Dermatol. 104:438-440(1995).
RN   [13]
RP   VARIANTS RDEB AND DDEB.
RX   MEDLINE=96220218; PubMed=8644729;
RA   Christiano A.M., McGrath J.A., Tan K.C., Uitto J.;
RT   "Glycine substitutions in the triple-helical region of type VII
RT   collagen result in a spectrum of dystrophic epidermolysis bullosa
RT   phenotypes and patterns of inheritance.";
RL   Am. J. Hum. Genet. 58:671-681(1996).
RN   [14]
RP   VARIANT RDEB ARG-2575.
RX   MEDLINE=96154068; PubMed=8592061;
RA   Shimizu H., McGrath J.A., Christiano A.M., Nishikawa T., Uitto J.;
RT   "Molecular basis of recessive dystrophic epidermolysis bullosa:
RT   genotype/phenotype correlation in a case of moderate clinical
RT   severity.";
```

```
RL   J. Invest. Dermatol. 106:119-124(1996).
RN   [15]
RP   VARIANT RDEB ARG-1782.
RX   MEDLINE=96183562; PubMed=8618018;
RA   Christiano A.M., McGrath J.A., Uitto J.;
RT   "Influence of the second COL7A1 mutation in determining the
RT   phenotypic severity of recessive dystrophic epidermolysis bullosa.";
RL   J. Invest. Dermatol. 106:766-770(1996).
RN   [16]
RP   VARIANT RDEB ASP-2073.
RX   MEDLINE=96310789; PubMed=8757758;
RA   Dunnill M.G.S., McGrath J.A., Richards A.J., Christiano A.M.,
RA   Uitto J., Pope F.M., Eady R.A.J.;
RT   "Clinicopathological correlations of compound heterozygous COL7A1
RT   mutations in recessive dystrophic epidermolysis bullosa.";
RL   J. Invest. Dermatol. 107:171-177(1996).
RN   [17]
RP   VARIANTS RDEB W-1982; G-2008; A-2025; E-2049; G-2063; W-2063 AND
RP   R-2575.
RX   MEDLINE=97465605; PubMed=9326325;
RA   Hovnanian A., Rochat A., Bodemer C., Petit E., Rivers C.A., Prost C.,
RA   Fraitag S., Christiano A.M., Uitto J., Lathrop M., Barrandon Y.,
RA   de Prost Y.;
RT   "Characterization of 18 new mutations in COL7A1 in recessive
RT   dystrophic epidermolysis bullosa provides evidence for distinct
RT   molecular mechanisms underlying defective anchoring fibril
RT   formation.";
RL   Am. J. Hum. Genet. 61:599-610(1997).
RN   [18]
RP   VARIANT RDEB ARG-1652.
RX   MEDLINE=98106792; PubMed=9444387;
RA   Cserhalmi-Friedman P.B., Karpati S., Horvath A., Christiano A.M.;
RT   "Identification of a glycine substitution and a splice site mutation
RT   in the type VII collagen gene in a proband with mitis recessive
RT   dystrophic epidermolysis bullosa.";
RL   Arch. Dermatol. Res. 289:640-645(1997).
RN   [19]
RP   VARIANTS DEB ARG-2009 AND ARG-2043.
RX   MEDLINE=97358588; PubMed=9215684;
RA   Winberg J.-O., Hammami-Hauasli N., Nilssen O., Anton-Lamprecht I.,
RA   Naylor S.L., Kerbacher K., Zimmermann M., Krajci P.,
RA   Gedde-Dahl T. Jr., Bruckner-Tuderman L.;
RT   "Modulation of disease severity of dystrophic epidermolysis bullosa by
RT   a splice site mutation in combination with a missense mutation in the
RT   COL7A1 gene.";
RL   Hum. Mol. Genet. 6:1125-1135(1997).
RN   [20]
RP   VARIANTS DDEB ASP-1519; ASP-2006; GLU-2015 AND ARG-2034.
RX   MEDLINE=98334662; PubMed=9668111;
RA   Hammami-Hauasli N., Schumann H., Raghunath M., Kilgus O., Luethi U.,
RA   Luger T., Bruckner-Tuderman L.;
RT   "Some, but not all, glycine substitution mutations in COL7A1 result in
RT   intracellular accumulation of collagen VII, loss of anchoring
RT   fibrils, and skin blistering.";
RL   J. Biol. Chem. 273:19228-19234(1998).
RN   [21]
RP   VARIANTS DEB CYS-2008; ARG-2207 AND SER-2775.
RX   MEDLINE=98410969; PubMed=9740253;
RA   Kon A., Pulkkinen L., Ishida-Yamamoto A., Hashimoto I., Uitto J.;
RT   "Novel COL7A1 mutations in dystrophic forms of epidermolysis
RT   bullosa.";
RL   J. Invest. Dermatol. 111:534-537(1998).
RN   [22]
RP   VARIANT RDEB ARG-1347.
RX   MEDLINE=99019477; PubMed=9804332;
RA   Terracina M., Posteraro P., Schubert M., Sonego G., Atzori F.,
RA   Zambruno G., Bruckner-Tuderman L., Castiglia D.;
RT   "Compound heterozygosity for a recessive glycine substitution and a
RT   splice site mutation in the COL7A1 gene causes an unusually mild form
RT   of localized recessive dystrophic epidermolysis bullosa.";
RL   J. Invest. Dermatol. 111:744-750(1998).
RN   [23]
RP   VARIANTS DEB TRP-2034; VAL-2040; ARG-2043; ARG-2064 AND ASP-2713.
RX   MEDLINE=99072663; PubMed=9856843;

  Query Match             4.5%;  Score 11;  DB 1;  Length 2944;
  Best Local Similarity   100.0%;  Pred. No. 0.38;
  Matches   11;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         71 PGLPGPRGDPG 81
               ||||||||||| 
Db       1328 PGLPGPRGDPG 1338


RESULT 10
CAC2_HAECO
ID   CAC2_HAECO     STANDARD;      PRT;   210 AA.
AC   P16252;
DT   01-AUG-1990 (Rel. 15, Created)
DT   01-AUG-1990 (Rel. 15, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Cuticle collagen 2C (Fragment).
GN   2C.
OS   Haemonchus contortus (Barber pole worm).
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Strongylida;
OC   Trichostrongyloidea; Haemonchidae; Haemonchinae; Haemonchus.
OX   NCBI_TaxID=6289;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90136718; PubMed=2615789;
RA   Shamansky L.M., Pratt D., Boisvenue R.J., Cox G.N.;
RT   "Cuticle collagen genes of Haemonchus contortus and Caenorhabditis
RT   elegans are highly conserved.";
RL   Mol. Biochem. Parasitol. 37:73-86(1989).
CC   -!- FUNCTION: NEMATODE CUTICLES ARE COMPOSED LARGELY OF COLLAGEN-LIKE
CC       PROTEINS. THE CUTICLE FUNCTIONS BOTH AS AN EXOSKELETON AND AS A
CC       BARRIER TO PROTECT THE WORM FROM ITS ENVIRONMENT.
CC   -!- MISCELLANEOUS: THIS PROTEIN SHOWS 4 POTENTIAL TRIPLE-HELICAL
CC       REGIONS, WHICH CONTAIN GLYCINE AS EVERY THIRD AMINO ACID.
CC   -!- MISCELLANEOUS: IN ALL NEMATODE CUTICLE COLLAGENS, THE POLYPEPTIDE
CC       CHAINS ARE COMPLEXED WITHIN THE CUTICLE BY DISULFIDE BONDS AND
CC       OTHER TYPES OF COVALENT CROSS-LINKS.
CC   -!- SIMILARITY: BELONGS TO THE CUTICULAR COLLAGEN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; J04670; AAA29172.1; -.
DR   InterPro; IPR000087; Collagen.
DR   Pfam; PF01391; Collagen; 3.
KW   Cuticle; Connective tissue; Repeat; Multigene family; Collagen.
FT   NON_TER       1       1
SQ   SEQUENCE   210 AA;  19562 MW;  E15FAA9A2DF3D74B CRC64;

  Query Match             4.1%;  Score 10;  DB 1;  Length 210;
  Best Local Similarity   100.0%;  Pred. No. 0.31;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         57 GAPGAPGEKG 66
              ||||||||||
Db        151 GAPGAPGEKG 160


RESULT 11
PSPA_CAVPO
ID   PSPA_CAVPO     STANDARD;      PRT;   247 AA.
AC   P50403;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Pulmonary surfactant-associated protein A precursor (SP-A) (PSP-A)
```

```
DE   (PSAP).
GN   SFTPA1 OR SFTPA OR SFTP1.
OS   Cavia porcellus (Guinea pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.
OX   NCBI_TaxID=10141;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Hartley; TISSUE=Lung;
RX   MEDLINE=98018900; PubMed=9357868;
RA   Yuan H.T., Gowan S., Kelly F.J., Bingle C.D.;
RT   "Cloning of guinea pig surfactant protein A defines a distinct
RT   cellular distribution pattern within the lung.";
RL   Am. J. Physiol. 273:L900-L906(1997).
CC   -!- FUNCTION: IN PRESENCE OF CALCIUM IONS, PSAP BINDS TO SURFACTANT
CC       PHOSPHOLIPIDS AND CONTRIBUTES TO LOWER THE SURFACE TENSION AT THE
CC       AIR-LIQUID INTERFACE IN THE ALVEOLI OF THE MAMMALIAN LUNG AND IS
CC       ESSENTIAL FOR NORMAL RESPIRATION.
CC   -!- SUBUNIT: OLIGOMERIC COMPLEX OF 6 SET OF HOMOTRIMERS.
CC   -!- SUBCELLULAR LOCATION: Extracellular.
CC   -!- MISCELLANEOUS: PULMONARY SURFACTANT CONSISTS OF 90% LIPID AND 10%
CC       PROTEIN. THERE ARE 4 SURFACTANT ASSOCIATED PROTEIN: 2 COLLAGENOUS,
CC       CARBOHYDRATE-BINDING GLYCOPROTEINS (SP-A AND SP-D) AND 2 SMALL
CC       HYDROPHOBIC PROTEINS (SP-B AND SP-C).
CC   -!- SIMILARITY: CONTAINS A COLLAGENOUS DOMAIN.
CC   -!- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U40869; AAB82952.1; -.
DR   HSSP; P22897; 1EGG.
DR   InterPro; IPR000087; Collagen.
DR   InterPro; IPR001304; Lectin_C.
DR   Pfam; PF00059; lectin_c; 1.
DR   Pfam; PF01391; Collagen; 1.
DR   SMART; SM00034; CLECT; 1.
DR   PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
KW   Glycoprotein; Calcium; Surface film; Gaseous exchange; Hydroxylation;
KW   Signal; Lectin; Collagen; Repeat.
FT   SIGNAL        1     19       POTENTIAL.
FT   CHAIN        20    247       PULMONARY SURFACTANT-ASSOCIATED PROTEIN
FT                                 A.
FT   DOMAIN       27     99       COLLAGEN-LIKE.
FT   DOMAIN      152    245       C-TYPE LECTIN (SHORT FORM).
FT   DISULFID    154    245       BY SIMILARITY.
FT   DISULFID    223    237       BY SIMILARITY.
FT   CARBOHYD     20     20       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    206    206       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   247 AA;  26104 MW;  D1BC86270EEFC932 CRC64;

  Query Match             4.1%;  Score 10;  DB 1;  Length 247;
  Best Local Similarity   100.0%;  Pred. No. 0.35;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         45 GLPGRDGRDG 54
              ||||||||||
Db         39 GLPGRDGRDG 48


RESULT 12
PSPA_RABIT
ID   PSPA_RABIT     STANDARD;      PRT;   247 AA.
AC   P12842;
DT   01-OCT-1989 (Rel. 12, Created)
DT   01-OCT-1989 (Rel. 12, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Pulmonary surfactant-associated protein A precursor (SP-A) (PSP-A)
DE   (PSAP).
GN   SFTPA1 OR SFTPA OR SFTP1.
OS   Oryctolagus cuniculus (Rabbit).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX   NCBI_TaxID=9986;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=88139348; PubMed=2830270;
RA   Boggaram V., Qing K., Mendelson C.R.;
RT   "The major apoprotein of rabbit pulmonary surfactant. Elucidation of
RT   primary sequence and cyclic AMP and developmental regulation.";
RL   J. Biol. Chem. 263:2939-2947(1988).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=New Zealand white; TISSUE=Liver;
RX   MEDLINE=92312742; PubMed=1616051;
RA   Chen Q., Boggaram V., Mendelson C.R.;
RT   "Rabbit lung surfactant protein A gene: identification of a lung-
RT   specific DNase I hypersensitive site.";
RL   Am. J. Physiol. 262:L662-L671(1992).
CC   -!- FUNCTION: IN PRESENCE OF CALCIUM IONS, PSAP BINDS TO SURFACTANT
CC       PHOSPHOLIPIDS AND CONTRIBUTES TO LOWER THE SURFACE TENSION AT THE
CC       AIR-LIQUID INTERFACE IN THE ALVEOLI OF THE MAMMALIAN LUNG AND IS
CC       ESSENTIAL FOR NORMAL RESPIRATION.
CC   -!- SUBUNIT: OLIGOMERIC COMPLEX OF 6 SET OF HOMOTRIMERS.
CC   -!- SUBCELLULAR LOCATION: Extracellular.
CC   -!- MISCELLANEOUS: PULMONARY SURFACTANT CONSISTS OF 90% LIPID AND 10%
CC       PROTEIN. THERE ARE 4 SURFACTANT ASSOCIATED PROTEIN: 2 COLLAGENOUS,
CC       CARBOHYDRATE-BINDING GLYCOPROTEINS (SP-A AND SP-D) AND 2 SMALL
CC       HYDROPHOBIC PROTEINS (SP-B AND SP-C).
CC   -!- SIMILARITY: CONTAINS A COLLAGENOUS DOMAIN.
CC   -!- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; J03542; AAA31465.1; -.
DR   EMBL; L19387; AAA31468.1; -.
DR   PIR; A29931; LNRBPS.
DR   HSSP; P22897; 1EGG.
DR   InterPro; IPR000087; Collagen.
DR   InterPro; IPR001304; Lectin_C.
DR   Pfam; PF00059; lectin_c; 1.
DR   Pfam; PF01391; Collagen; 2.
DR   SMART; SM00034; CLECT; 1.
DR   PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
KW   Glycoprotein; Calcium; Surface film; Gaseous exchange; Hydroxylation;
KW   Signal; Lectin; Collagen; Repeat.
FT   SIGNAL        1     15       POTENTIAL.
FT   CHAIN        16    247       PULMONARY SURFACTANT-ASSOCIATED PROTEIN
FT                                 A.
FT   DOMAIN       27     99       COLLAGEN-LIKE.
FT   DOMAIN      152    247       C-TYPE LECTIN (SHORT FORM).
FT   DISULFID    154    245       BY SIMILARITY.
FT   DISULFID    223    237       BY SIMILARITY.
FT   CARBOHYD    206    206       N-LINKED (GLCNAC. . .) (PROBABLE).
FT   VARIANT      12     12       S -> P.
FT   CONFLICT     57     60       GPMG -> APWA (IN REF. 2).
SQ   SEQUENCE   247 AA;  26071 MW;  289634054CBC8CB4 CRC64;

  Query Match             4.1%;  Score 10;  DB 1;  Length 247;
  Best Local Similarity   100.0%;  Pred. No. 0.35;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         45 GLPGRDGRDG 54
```

```
                  ||||||||||
Db          39 GLPGRDGRDG 48


RESULT 13
PSPA_CANFA
ID   PSPA_CANFA     STANDARD;      PRT;   248 AA.
AC   P06908;
DT   01-JAN-1988 (Rel. 06, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Pulmonary surfactant-associated protein A precursor (SP-A) (PSP-A)
DE   (PSAP).
GN   SFTPA1 OR SFTPA OR SFTP1.
OS   Canis familiaris (Dog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX   NCBI_TaxID=9615;
RN   [1]
RP   SEQUENCE FROM N.A., AND SIGNAL SEQUENCE CLEAVAGE SITE.
RX   MEDLINE=86016705; PubMed=3863100;
RA   Benson B., Hawgood S., Schilling J., Clements J., Damm D., Cordell B.,
RA   White R.T.;
RT   "Structure of canine pulmonary surfactant apoprotein: cDNA and
RT   complete amino acid sequence.";
RL   Proc. Natl. Acad. Sci. U.S.A. 82:6379-6383(1985).
RN   [2]
RP   C-TYPE LECTIN DOMAIN.
RX   MEDLINE=87115834; PubMed=3808053;
RA   Patthy L.;
RT   "Is lung surfactant protein a lectin-collagen hybrid?";
RL   Nature 325:490-490(1987).
CC   -!- FUNCTION: IN PRESENCE OF CALCIUM IONS, PSAP BINDS TO SURFACTANT
CC       PHOSPHOLIPIDS AND CONTRIBUTES TO LOWER THE SURFACE TENSION AT THE
CC       AIR-LIQUID INTERFACE IN THE ALVEOLI OF THE MAMMALIAN LUNG AND IS
CC       ESSENTIAL FOR NORMAL RESPIRATION.
CC   -!- SUBUNIT: OLIGOMERIC COMPLEX OF 6 SET OF HOMOTRIMERS.
CC   -!- SUBCELLULAR LOCATION: Extracellular.
CC   -!- MISCELLANEOUS: PULMONARY SURFACTANT CONSISTS OF 90% LIPID AND 10%
CC       PROTEIN. THERE ARE 4 SURFACTANT ASSOCIATED PROTEIN: 2 COLLAGENOUS,
CC       CARBOHYDRATE-BINDING GLYCOPROTEINS (SP-A AND SP-D) AND 2 SMALL
CC       HYDROPHOBIC PROTEINS (SP-B AND SP-C).
CC   -!- SIMILARITY: CONTAINS A COLLAGENOUS DOMAIN.
CC   -!- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M11769; AAA30887.1; -.
DR   PIR; A25296; LNDGPS.
DR   HSSP; P22897; 1EGG.
DR   InterPro; IPR000087; Collagen.
DR   InterPro; IPR001304; Lectin_C.
DR   Pfam; PF00059; lectin_c; 1.
DR   Pfam; PF01391; Collagen; 2.
DR   SMART; SM00034; CLECT; 1.
DR   PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
KW   Glycoprotein; Calcium; Surface film; Gaseous exchange; Hydroxylation;
KW   Signal; Lectin; Collagen; Repeat.
FT   SIGNAL       1     17
FT   CHAIN       18    248       PULMONARY SURFACTANT-ASSOCIATED PROTEIN
FT                                 A.
FT   DOMAIN       28    100       COLLAGEN-LIKE.
FT   DOMAIN      153    248       C-TYPE LECTIN (SHORT FORM).
FT   DISULFID    155    246       BY SIMILARITY.
FT   DISULFID    224    238       BY SIMILARITY.
FT   CARBOHYD     20     20       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    207    207       N-LINKED (GLCNAC. . .) (PROBABLE).
SQ   SEQUENCE   248 AA;  26268 MW;  340FE95D4E2502C0 CRC64;

  Query Match             4.1%;  Score 10;  DB 1;  Length 248;
  Best Local Similarity   100.0%;  Pred. No. 0.35;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          45 GLPGRDGRDG 54
               ||||||||||
Db          40 GLPGRDGRDG 49


RESULT 14
PSPA_HUMAN
ID   PSPA_HUMAN     STANDARD;      PRT;   248 AA.
AC   P07714;
DT   01-APR-1988 (Rel. 07, Created)
DT   01-APR-1990 (Rel. 14, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Pulmonary surfactant-associated protein A precursor (SP-A) (PSP-A)
DE   (PSAP) (Alveolar proteinosis protein) (35 kDa pulmonary surfactant-
DE   associated protein).
GN   SFTPA1 OR SFTPA OR SFTP1 OR PSAP.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=86250832; PubMed=3755136;
RA   Floros J., Steinbrink R., Jacobs K., Phelps D., Kriz R., Recny M.,
RA   Sultzman L., Jones S., Taeusch H.W., Frank H.A., Fritsch E.F.;
RT   "Isolation and characterization of cDNA clones for the 35-kDa
RT   pulmonary surfactant-associated protein.";
RL   J. Biol. Chem. 261:9029-9033(1986).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=86014366; PubMed=2995821;
RA   White R.T., Damm D., Miller J., Spratt K., Schilling J., Hawgood S.,
RA   Benson B., Cordell B.;
RT   "Isolation and characterization of the human pulmonary surfactant
RT   apoprotein gene.";
RL   Nature 317:361-363(1985).
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92198680; PubMed=1372511;
RA   Katyal S.L., Singh G., Locker J.L.;
RT   "Characterization of a second human pulmonary surfactant-associated
RT   protein SP-A gene.";
RL   Am. J. Respir. Cell Mol. Biol. 6:446-452(1992).
CC   -!- FUNCTION: IN PRESENCE OF CALCIUM IONS, PSAP BINDS TO SURFACTANT
CC       PHOSPHOLIPIDS AND CONTRIBUTES TO LOWER THE SURFACE TENSION AT THE
CC       AIR-LIQUID INTERFACE IN THE ALVEOLI OF THE MAMMALIAN LUNG AND IS
CC       ESSENTIAL FOR NORMAL RESPIRATION.
CC   -!- SUBUNIT: OLIGOMERIC COMPLEX OF 6 SET OF HOMOTRIMERS.
CC   -!- SUBCELLULAR LOCATION: Extracellular.
CC   -!- MISCELLANEOUS: PULMONARY SURFACTANT CONSISTS OF 90% LIPID AND 10%
CC       PROTEIN. THERE ARE 4 SURFACTANT ASSOCIATED PROTEIN: 2 COLLAGENOUS,
CC       CARBOHYDRATE-BINDING GLYCOPROTEINS (SP-A AND SP-D) AND 2 SMALL
CC       HYDROPHOBIC PROTEINS (SP-B AND SP-C).
CC   -!- SIMILARITY: CONTAINS A COLLAGENOUS DOMAIN.
CC   -!- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M13686; AAA60211.1; -.
DR   EMBL; K03475; AAA36520.1; -.
```

```
DR   EMBL; M30838; AAA36510.1; -.
DR   EMBL; M68519; AAA60319.1; -.
DR   PIR; A24622; LNHUPS.
DR   PIR; A25720; LNHUP6.
DR   PIR; B25720; LNHUP1.
DR   HSSP; P22897; 1EGG.
DR   Genew; HGNC:10799; SFTPA2.
DR   Genew; HGNC:10798; SFTPA1.
DR   MIM; 178630; -.
DR   InterPro; IPR000087; Collagen.
DR   InterPro; IPR001304; Lectin_C.
DR   Pfam; PF00059; lectin_c; 1.
DR   Pfam; PF01391; Collagen; 2.
DR   SMART; SM00034; CLECT; 1.
DR   PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
KW   Glycoprotein; Calcium; Surface film; Gaseous exchange; Hydroxylation;
KW   Signal; Lectin; Collagen; Repeat; Polymorphism.
FT   SIGNAL        1     20
FT   CHAIN        21    248       PULMONARY SURFACTANT-ASSOCIATED PROTEIN
FT                                 A.
FT   DOMAIN       28    100       COLLAGEN-LIKE.
FT   DOMAIN      153    248       C-TYPE LECTIN (SHORT FORM).
FT   DISULFID    155    246       BY SIMILARITY.
FT   DISULFID    224    238       BY SIMILARITY.
FT   CARBOHYD    207    207       N-LINKED (GLCNAC. . .) (PROBABLE).
FT   VARIANT       9      9       N -> T.
FT                                 /FTId=VAR_004184.
FT   VARIANT      50     50       L -> V.
FT                                 /FTId=VAR_012231.
FT   VARIANT      66     66       M -> T.
FT                                 /FTId=VAR_004185.
FT   VARIANT      73     73       D -> N.
FT                                 /FTId=VAR_004186.
FT   VARIANT      81     81       I -> V.
FT                                 /FTId=VAR_004187.
FT   VARIANT      85     85       C -> R.
FT                                 /FTId=VAR_004188.
FT   VARIANT     219    219       R -> W.
FT                                 /FTId=VAR_012232.
FT   VARIANT     223    223       Q -> K.
FT                                 /FTId=VAR_012233.
FT   CONFLICT     19     19       A -> V (IN REF. 2).
FT   CONFLICT     45     45       D -> H (IN REF. 2).
FT   CONFLICT     54     54       P -> L (IN REF. 2).
FT   CONFLICT     91     91       P -> A (IN REF. 3).
FT   CONFLICT    100    100       P -> R (IN REF. 2; AAA36510).
FT   CONFLICT    247    247       E -> D (IN REF. 3).
SQ   SEQUENCE   248 AA;  26214 MW;  6A9F0C3488BF3633 CRC64;

  Query Match             4.1%;  Score 10;  DB 1;  Length 248;
  Best Local Similarity   100.0%;  Pred. No. 0.35;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         45 GLPGRDGRDG 54
              ||||||||||
Db         40 GLPGRDGRDG 49


RESULT 15
PSPA_MOUSE
ID   PSPA_MOUSE     STANDARD;      PRT;   248 AA.
AC   P35242;
DT   01-FEB-1994 (Rel. 28, Created)
DT   01-FEB-1994 (Rel. 28, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Pulmonary surfactant-associated protein A precursor (SP-A) (PSP-A)
DE   (PSAP).
GN   SFTPA1 OR SFTPA OR SFTP1 OR SFTP-1.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=DBA/2J;
RX   MEDLINE=93072386; PubMed=1443158;
RA   Korfhagen T.R., Bruno M.D., Glasser S.W., Ciraolo P.J., Whitsett J.A.,
RA   Lattier D.L., Wikenheiser K.A., Clark J.C.;
RT   "Murine pulmonary surfactant SP-A gene: cloning, sequence, and
RT   transcriptional activity.";
RL   Am. J. Physiol. 263:L546-L554(1992).
CC   -!- FUNCTION: IN PRESENCE OF CALCIUM IONS, PSAP BINDS TO SURFACTANT
CC       PHOSPHOLIPIDS AND CONTRIBUTES TO LOWER THE SURFACE TENSION AT THE
CC       AIR-LIQUID INTERFACE IN THE ALVEOLI OF THE MAMMALIAN LUNG AND IS
CC       ESSENTIAL FOR NORMAL RESPIRATION.
CC   -!- SUBUNIT: OLIGOMERIC COMPLEX OF 6 SET OF HOMOTRIMERS.
CC   -!- SUBCELLULAR LOCATION: Extracellular.
CC   -!- MISCELLANEOUS: PULMONARY SURFACTANT CONSISTS OF 90% LIPID AND 10%
CC       PROTEIN. THERE ARE 4 SURFACTANT ASSOCIATED PROTEIN: 2 COLLAGENOUS,
CC       CARBOHYDRATE-BINDING GLYCOPROTEINS (SP-A AND SP-D) AND 2 SMALL
CC       HYDROPHOBIC PROTEINS (SP-B AND SP-C).
CC   -!- SIMILARITY: CONTAINS A COLLAGENOUS DOMAIN.
CC   -!- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; S48768; AAB24274.1; -.
DR   PIR; A48853; A48853.
DR   HSSP; P35247; 1B08.
DR   MGD; MGI:109518; Sftpa.
DR   InterPro; IPR000087; Collagen.
DR   InterPro; IPR001304; Lectin_C.
DR   Pfam; PF00059; lectin_c; 1.
DR   Pfam; PF01391; Collagen; 2.
DR   SMART; SM00034; CLECT; 1.
DR   PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
KW   Glycoprotein; Calcium; Surface film; Gaseous exchange; Hydroxylation;
KW   Signal; Lectin; Collagen; Repeat.
FT   SIGNAL        1     20       BY SIMILARITY.
FT   CHAIN        21    248       PULMONARY SURFACTANT-ASSOCIATED PROTEIN
FT                                 A.
FT   DOMAIN       28    100       COLLAGEN-LIKE.
FT   DOMAIN      153    248       C-TYPE LECTIN (SHORT FORM).
FT   DISULFID    155    246       BY SIMILARITY.
FT   DISULFID    224    238       BY SIMILARITY.
FT   CARBOHYD     21     21       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    207    207       N-LINKED (GLCNAC. . .) (PROBABLE).
SQ   SEQUENCE   248 AA;  26157 MW;  6688BF070E3EB9AE CRC64;

  Query Match             4.1%;  Score 10;  DB 1;  Length 248;
  Best Local Similarity   100.0%;  Pred. No. 0.35;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         45 GLPGRDGRDG 54
              ||||||||||
Db         40 GLPGRDGRDG 49


Search completed: June 18, 2003, 15:04:24
Job time : 32 secs
```

GenCore version 5.1.6


OM protein - protein search, using sw model

Run on: June 18, 2003, 14:48:42 ; Search time 39 Seconds
(without alignments)
598.991 Million cell updates/sec

Title: US-09-943-851A-42
Perfect score: 1325
Sequence: 1 MRPLLVLLLLGLAAGSPPLD.....DSTFSGFLVYSDWHSSPVFA 243

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_73:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 1207 | 91.1 | 219 | 2 | T14782 | hypothetical prote |
| 2 | 424 | 32.0 | 244 | 2 | JC4708 | gelatin-binding 28 |
| 3 | 418.5 | 31.6 | 680 | 1 | CGHU1D | collagen alpha 1(X |
| 4 | 416.5 | 31.4 | 674 | 2 | S23297 | collagen alpha 1(X |
| 5 | 412 | 31.1 | 246 | 2 | S29328 | complement subcomp |
| 6 | 411.5 | 31.1 | 674 | 2 | S13301 | collagen alpha 1(X |
| 7 | 408 | 30.8 | 680 | 2 | S31216 | collagen alpha 1(X |
| 8 | 399 | 30.1 | 245 | 1 | C1HUQC | complement subcomp |
| 9 | 374 | 28.2 | 744 | 1 | A34246 | collagen alpha 1(V |
| 10 | 370 | 27.9 | 744 | 1 | S23298 | collagen alpha 1(V |
| 11 | 363.5 | 27.4 | 743 | 1 | S23779 | collagen alpha 1(V |
| 12 | 363 | 27.4 | 635 | 2 | A57131 | collagen alpha 2(V |
| 13 | 362 | 27.3 | 253 | 1 | C1HUQB | complement subcomp |
| 14 | 360.5 | 27.2 | 253 | 2 | S49158 | complement protein |
| 15 | 359 | 27.1 | 744 | 2 | S15435 | collagen alpha 1(V |
| 16 | 351 | 26.5 | 253 | 2 | I49560 | complement C1q B c |
| 17 | 316 | 23.8 | 423 | 2 | A55797 | collagen precursor |
| 18 | 314 | 23.7 | 245 | 1 | C1HUQA | complement subcomp |
| 19 | 314 | 23.7 | 245 | 2 | S19018 | complement subcomp |
| 20 | 280.5 | 21.2 | 215 | 2 | B48150 | hibernation-relate |
| 21 | 277.5 | 20.9 | 215 | 2 | C48150 | hibernation-relate |
| 22 | 261.5 | 19.7 | 196 | 2 | A48150 | hibernation-relate |
| 23 | 238.5 | 18.0 | 992 | 2 | T08772 | hypothetical prote |
| 24 | 230.5 | 17.4 | 1049 | 1 | CGBO7S | collagen alpha 1(I |
| 25 | 227.5 | 17.2 | 248 | 2 | I51921 | pulmonary surfacta |
| 26 | 225.5 | 17.0 | 248 | 1 | LNHUP1 | pulmonary surfacta |
| 27 | 225 | 17.0 | 325 | 2 | T32248 | hypothetical prote |
| 28 | 224 | 16.9 | 636 | 2 | S41067 | collagen alpha 1(I |
| 29 | 223 | 16.8 | 1464 | 2 | S59856 | collagen alpha 1(I |
| 30 | 222.5 | 16.8 | 1758 | 2 | T29350 | hypothetical prote |
| 31 | 222.5 | 16.8 | 1759 | 2 | T29351 | collagen alpha 2(I |
| 32 | 222 | 16.8 | 360 | 2 | T37285 | collagen dpy-2 - C |
| 33 | 221.5 | 16.7 | 170 | 2 | B57131 | collagen alpha 2(V |
| 34 | 221 | 16.7 | 886 | 2 | I50694 | collagen alpha 1(I |
| 35 | 221 | 16.7 | 1019 | 1 | A32856 | collagen alpha 1(V |
| 36 | 220 | 16.6 | 248 | 1 | LNDGPS | pulmonary surfacta |
| 37 | 220 | 16.6 | 341 | 2 | T16296 | hypothetical prote |
| 38 | 220 | 16.6 | 380 | 2 | T28888 | cuticle collagen d |
| 39 | 219.5 | 16.6 | 247 | 1 | LNRBPS | pulmonary surfacta |
| 40 | 219.5 | 16.6 | 381 | 2 | T27806 | hypothetical prote |
| 41 | 219.5 | 16.6 | 673 | 1 | CGBO6C | collagen alpha 1(I |
| 42 | 219.5 | 16.6 | 1466 | 1 | CGHU7L | collagen alpha 1(I |
| 43 | 219 | 16.5 | 671 | 1 | CGRT1S | collagen alpha 1(I |
| 44 | 218.5 | 16.5 | 1453 | 2 | S21626 | collagen alpha 1(I |
| 45 | 217 | 16.4 | 283 | 2 | T29980 | hypothetical prote |

ALIGNMENTS

RESULT 1
T14782
hypothetical protein DKFZp586B0621.1 - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 02-Sep-2000
C;Accession: T14782
R;Ottenwaelder, B.; Obermaier, B.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, August 1999
A;Reference number: Z18184
A;Accession: T14782
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-219 <OTT>
A;Cross-references: EMBL:AL110261
A;Experimental source: adult uterus; clone DKFZp586B0621
C;Genetics:
A;Note: DKFZp586B0621.1
C;Superfamily: complement C1q carboxyl-terminal homology

Query Match 91.1%; Score 1207; DB 2; Length 219;
Best Local Similarity 100.0%; Pred. No. 1.3e-77;
Matches 219; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          25 PSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPGAPGEKGEGGRPGLPGPRGDPGPRG 84
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           1 PSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPGAPGEKGEGGRPGLPGPRGDPGPRG 60

Qy          85 EAGPAGPTGPAGECSVPPRSAFSAKRSESRVPPPSDAPLPFDRVLVNEQGHYDAVTGKFT 144
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          61 EAGPAGPTGPAGECSVPPRSAFSAKRSESRVPPPSDAPLPFDRVLVNEQGHYDAVTGKFT 120

Qy         145 CQVPGVYYFAVHATVYRASLQFDLVKNGESIASFFQFFGGWPKPASLSGGAMVRLEPEDQ 204
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         121 CQVPGVYYFAVHATVYRASLQFDLVKNGESIASFFQFFGGWPKPASLSGGAMVRLEPEDQ 180

Qy         205 VWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWHSSPVFA 243
               |||||||||||||||||||||||||||||||||||||||
Db         181 VWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWHSSPVFA 219
```

RESULT 2
JC4708
gelatin-binding 28K protein precursor - human
N;Alternate names: adipose specific collagen-like factor
C;Species: Homo sapiens (man)
C;Date: 10-May-1996 #sequence_revision 19-Jul-1996 #text_change 20-Sep-1999
C;Accession: JC4708; JC4944
R;Maeda, K.; Okubo, K.; Shimomura, I.; Funahashi, T.; Matsuzawa, Y.; Matsubara, K.
Biochem. Biophys. Res. Commun. 221, 286-289, 1996
A;Title: cDNA cloning and expression of a novel adipose specific collagen-like factor
A;Reference number: JC4708; MUID:96224171; PMID:8619847
A;Accession: JC4708

```
A;Molecule type: mRNA
A;Residues: 1-244 <MAE>
A;Cross-references: DDBJ:D45371; NID:g871886; PIDN:BAA08227.1; PID:g871887
A;Experimental source: adipose tissue
R;Nakano, Y.; Tobe, T.; Choi-Miura, N.H.; Mazda, T.; Tomita, M.
J. Biochem. 120, 803-812, 1996
A;Title: Isolation and characterization of GBP28, a novel gelatin-binding protein purifi
A;Reference number: JC4944; MUID:97103474; PMID:8947845
A;Accession: JC4944
A;Molecule type: protein
A;Residues: 19-38;93-100;101-112;135-149;173-178 <NAK>
C;Comment: This protein is an endogenous factor that binds with a collagen-like domain.
C;Genetics:
A;Gene: apM1
C;Superfamily: unassigned collagens; complement C1q carboxyl-terminal homology
C;Keywords: adipose tissue; glycoprotein; hydroxyproline
F;1-18/Domain: signal sequence #status predicted <SIG>
F;19-244/Product: gelatin-binding 28kDa protein #status experimental <MAT>
F;42-107/Region: collagen-like
F;114-241/Domain: complement C1q carboxyl-terminal homology <C1Q>
F;95/Modified site: 4-hydroxyproline (Pro) #status experimental
F;230/Binding site: carbohydrate (Asn) (covalent) #status predicted

  Query Match             32.0%;  Score 424;  DB 2;  Length 244;
  Best Local Similarity   40.5%;  Pred. No. 7.7e-23;
  Matches  106;  Conservative   26;  Mismatches   72;  Indels   58;  Gaps   10;

Qy          6 VLLLLGLAAGSPPLDDNKIPSLCPGH--------PG-------------LPGTPGHHGSQ 44
              ||||| |                |||         ||             : | ||| |
Db          7 VLLLLAL----------------PGHDQETTTQGPGVLLPLPKGACTGWMAGIPGHPGHN 50

Qy         45 GLPGRDGRDGRDGAPGAPGEKGEGGRPGLPGPRGD---------PGPRGEAGPAGPTGPA 95
              | |||||||      | |||||| | ||| ||:||          ||||  |  |  |
Db         51 GAPGRDGRD------GTPGEKGEKGDPGLIGPKGDIGETGVPGAEGPRGFPGIQGRKGEP 104

Qy         96 GECSVPPRSAFSAKRSESRVPPPSDAPLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAV 155
              || :  |||||    |: | | : |: | ::  |:| |||  |||| | :||:||||
Db        105 GEGAYVYRSAFSV-GLETYVTIP-NMPIRFTKIFYNQQNHYDGSTGKFHCNIPGLYYFAY 162

Qy        156 HATVYRASLQFDLVKNGES-IASFFQFFGGWPKPASLSGGAMVRLEPEDQVWVQV-GVGD 213
              | |||   ::  | |  :: : :: |:       |  ||  :: ||  ||||:|| | |:
Db        163 HITVYMKDVKVSLFKKDKAMLFTYDQYQENNVDQA--SGSVLLHLEVGDQVWLQVYGEGE 220

Qy        214 YIGIYASIKTDSTFSGFLVYSD 235
                |:||    ||||:|||:| |
Db        221 RNGLYADNDNDSTFTGFLLYHD 242


RESULT 3
CGHU1D
collagen alpha 1(X) chain precursor - human
N;Alternate names: procollagen alpha 1(X) chain
C;Species: Homo sapiens (man)
C;Date: 22-Nov-1993 #sequence_revision 03-Nov-1995 #text_change 22-Jun-1999
C;Accession: S26396; S30086; S15826; S18249; A43901; I51870; S21856
R;Reichenberger, E.; Beier, F.; LuValle, P.; Olsen, B.R.; von der Mark, K.; Bertling, W.
FEBS Lett. 311, 305-310, 1992
A;Title: Genomic organization and full-length cDNA sequence of human collagen X.
A;Reference number: S26396; MUID:93012005; PMID:1397333
A;Accession: S26396
A;Molecule type: DNA
A;Residues: 1-680 <REI>
A;Cross-references: EMBL:X68952; EMBL:X72578; EMBL:X72579; EMBL:X72580; GB:S47714; GB:S4
R;Apte, S.S.
submitted to the EMBL Data Library, March 1992
A;Reference number: S30085
A;Accession: S30086
A;Molecule type: DNA
A;Residues: 'TIPFYGWVCWVCLL',52-680 <APT>
A;Cross-references: EMBL:X65120; NID:g23129
A;Note: the initial difference is probably due to translation of an intronic sequence
R;Apte, S.; Mattei, M.G.; Olsen, B.R.
FEBS Lett. 282, 393-396, 1991
A;Title: Cloning of human alpha-1(X) collagen DNA and localization of the COL10A1 gen
A;Reference number: S15826; MUID:91243838; PMID:2037056
A;Accession: S15826
A;Molecule type: DNA
A;Residues: 561-647,'G',649-666 <AP2>
A;Cross-references: EMBL:X58879; NID:g30013; PIDN:CAA41686.1; PID:g30014
R;Thomas, J.T.; Cresswell, C.J.; Rash, B.; Nicolai, H.; Jones, T.; Solomon, E.; Grant
Biochem. J. 280, 617-623, 1991
A;Title: The human collagen X gene. Complete primary translated sequence and chromoso
A;Reference number: S18249; MUID:92109659; PMID:1764025
A;Accession: S18249
A;Molecule type: DNA
A;Residues: 1-26,'T',28-680 <THO>
A;Cross-references: EMBL:X60382; NID:g30094; PIDN:CAA42933.1; PID:g30095
A;Note: the sequence from Fig. 3 is inconsistent with that from Fig. 2 in having 532-
R;Reichenberger, E.; Aigner, T.; von der Mark, K.; Stoss, H.; Bertling, W.
Dev. Biol. 148, 562-572, 1991
A;Title: In situ hybridization studies on the expression of type X collagen in fetal
A;Reference number: A43901; MUID:92077285; PMID:1743401
A;Accession: A43901
A;Molecule type: mRNA
A;Residues: 547-656 <RE2>
A;Cross-references: GB:M74050; GB:D57494; NID:g339884; PIDN:AAA61221.1; PID:g553796
A;Note: sequence extracted from NCBI backbone (NCBIN:69012, NCBIP:69014)
R;Wallis, G.A.; Rash, B.; Sweetman, W.A.; Thomas, J.T.; Super, M.; Evans, G.; Grant,
Am. J. Hum. Genet. 54, 169-178, 1994
A;Title: Amino acid substitutions of conserved residues in the carboxyl-terminal doma
pe Schmid.
A;Reference number: I51870; MUID:94136476; PMID:8304336
A;Accession: I51870
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 520-597,'D',599-680 <WAL>
A;Cross-references: GB:S68531; NID:g545180; PIDN:AAC60615.1; PID:g545181
A;Note: mutant sequence from patient with metaphyseal chondrodysplasia type Schmid
A;Note: a second mutant sequence with 614-Pro is also described
C;Comment: Prolines and lysines at the third position of the tripeptide repeating uni
ed and subsequently O-glycosylated.
C;Genetics:
A;Gene: GDB:COL10A1
A;Cross-references: GDB:128635; OMIM:120110
A;Map position: 6q21-6q22
A;Introns: 52/1
A;Note: a defect in this gene may cause Schmid metaphyseal chondrodysplasia
C;Complex: type X collagen may be a homotrimer
C;Function:
A;Description: structural component of extracellular fibrous polymer specifically and
be important for skeletogenesis
C;Superfamily: collagen alpha 1(VIII) chain; complement C1q carboxyl-terminal homolog
C;Keywords: coiled coil; extracellular matrix; glycoprotein; homotrimer; hydroxylysin
F;1-18/Domain: signal sequence #status predicted <SIG>
F;19-680/Product: collagen alpha 1(X) chain #status predicted <MAT>
F;19-56/Domain: amino-terminal nonhelical #status predicted <NC2>
F;57-519/Region: interrupted helical
F;520-680/Domain: amino-terminal nonhelical #status predicted <NC1>
F;553-679/Domain: complement C1q carboxyl-terminal homology <C1Q>
F;617/Binding site: carbohydrate (Asn) (covalent) #status predicted

  Query Match             31.6%;  Score 418.5;  DB 1;  Length 680;
  Best Local Similarity   34.9%;  Pred. No. 5.4e-22;
  Matches   98;  Conservative   34;  Mismatches   82;  Indels   67;  Gaps    6;

Qy         15 GSPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPGAPGEKGEGGRPGLP 74
              |:| |  |      | |||||  |  |::|:|| :|  |  |||| || :|  | ||:|
Db        401 GNPGLPGPKGDPGVGGPPGLPGPVGPAGAKGMPGHNGEAGPRGAPGIPGTRGPIGPPGIP 460

Qy         75 ----------------------------GPRGDPGPRGEA------GPAGPTGPAGECS 99
                                          || | |||||  :      || || || |:
Db        461 GFPGSKGDPGSPGPPGPAGIATKGLNGPTGPPGPPGPRGHSGEPGLPGPPGPPGPPGQAV 520

Qy        100 VP--------------------------PRSAFSAKRSESRVPPPSDAPLPFDRVLVNE 132
              :|                          | |||:   |::   |    |:|||::| |
Db        521 MPEGFIKAGQRPSLSGTPLVSANQGVTGMPVSAFTVILSKAY--PAIGTPIPFDKILYNR 578
```

```
Qy        133 QGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGESIA-SFFQFFGGWPKPASL 191
              | |||  || ||||:||:|||: |  |    :   | |||  :  :: ::  |:   |
Db        579 QQHYDPRTGIFTCQIPGIYYFSYHVHVKGTHVWVGLYKNGTPVMYTYDEYTKGYLDQA-- 636

Qy        192 SGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLV 232
              || |:: |   ||||:|:   :  |:|:|    |:||||||
Db        637 SGSAIIDLTENDQVWLQLPNAESNGLYSSEYVHSSFSGFLV 677


RESULT 4
S23297
collagen alpha 1(X) chain precursor - chicken
N;Alternate names: type X collagen
C;Species: Gallus gallus (chicken)
C;Date: 07-Oct-1994 #sequence_revision 10-Nov-1995 #text_change 13-Aug-1999
C;Accession: S23297; A31896; S65594; S77711; I50218
R;Ninomiya, Y.; Castagnola, P.; Gerecke, D.; Gordon, M.K.; Jacenko, O.; LuValle, P.; McC
maguchi, N.; Olsen, B.R.
in Extracellular Matrix Genes, Sandell L.J. and Boyd C.D., eds., pp.79-114, Academic Pre
A;Title: The molecular biology of collagens with short triple-helical domains.
A;Reference number: S22243
A;Accession: S23297
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-674 <NIN>
R;LuValle, P.; Ninomiya, Y.; Rosenblum, N.D.; Olsen, B.R.
J. Biol. Chem. 263, 18378-18385, 1988
A;Title: The type X collagen gene. Intron sequences split the 5'-untranslated region and
A;Reference number: A31896; MUID:89054019; PMID:2461368
A;Accession: A31896
A;Molecule type: mRNA
A;Residues: 1-75 <LUV>
R;Ninomiya, Y.; Gordon, M.; van der Rest, M.; Schmid, T.; Linsenmayer, T.; Olsen, B.R.
J. Biol. Chem. 261, 5041-5050, 1986
A;Title: The developmentally regulated type X collagen gene contains a long open reading
A;Reference number: I50218; MUID:86168227; PMID:3082876
A;Accession: S65594
A;Molecule type: DNA
A;Residues: 'T',9,'D',11-12,'EDQMKLYILFTM',30-31,'TCKSGRAFTTYMILQNVMADLVSSHT',48-89,'L',
629,'PQAVLSLISWRTIKCGSSCQIQNPMVSIPLNMFILLSQVSYLLKSNNIPLTMS' <NIN1>
A;Cross-references: EMBL:M13496; NID:g211699; PIDN:AAA48736.1; PID:g211700
A;Accession: S77711
A;Molecule type: protein
A;Residues: 104-112,'X',114-117;453-466 <NIN2>
C;Superfamily: collagen alpha 1(VIII) chain; complement C1q carboxyl-terminal homology
C;Keywords: coiled coil; extracellular matrix; glycoprotein; homotrimer; hydroxyproline;
F;1-18/Domain: signal sequence #status predicted <SIG>
F;547-673/Domain: complement C1q carboxyl-terminal homology <C1Q>
F;453,456/Modified site: hydroxyproline (Pro) #status experimental
F;611/Binding site: carbohydrate (Asn) (covalent) #status predicted

  Query Match             31.4%;  Score 416.5;  DB 2;  Length 674;
  Best Local Similarity   36.0%;  Pred. No. 7.3e-22;
  Matches   96;  Conservative   34;  Mismatches   84;  Indels   53;  Gaps    6;

Qy         14 AGSPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPGAPGEKGEGGRPGL 73
              || | |     |    | ||: | ||  |  |:||  |  |  | ||||| ||| | |||
Db        408 AGPPGLPGPVGPQGVKGVPGINGEPGPRGPSGIPGIRGPIGPPGMPGAPGAKGEAGAPGL 467

Qy         74 PGPR--------------GDPGPRGEA------GPAGPTGPAGECSVP------------ 101
              |||              | |||:| :       || || || |: ::|
Db        468 PGPAGIATKGLRGPMGPPGPPGPKGNSGEPGLPGPPGPPGPPGQSTIPEGYVKGESRELS 527

Qy        102 ---------------PRSAFSAKRSESRVPPPSDAPLPFDRVLVNEQGHYDAVTGKFTC 145
                            | |||:   |::   | :  |: ||::| | | |||  || |||
Db        528 GMSFMKAGANQALTGMPVSAFTVILSKAY--PGATVPIKFDKILYNRQQHYDPRTGIFTC 585

Qy        146 QVPGVYYFAVHATVYRASLQFDLVKNGESIA-SFFQFFGGWPKPASLSGGAMVRLEPEDQ 204
              ::||:|||: |      ::   | |||  :  :: ::  |:   |  || |:: |   ||
Db        586 RIPGLYYFSYHVHAKGTNVWVALYKNGSPVMYTYDEYQKGYLDQA--SGSAVIDLMENDQ 643

Qy        205 VWVQVGVGDYIGIYASIKTDSTFSGFL 231
              ||:|:   :  |:|:|    |:||||
Db        644 VWLQLPNSESNGLYSSEYVHSSFSGFL 670


RESULT 5
S29328
complement subcomponent C1q chain C - mouse
C;Species: Mus musculus (house mouse)
C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 20-Aug-1999
C;Accession: S29328
R;Petry, F.; Reid, K.B.M.; Loos, M.
Eur. J. Biochem. 209, 129-134, 1992
A;Title: Isolation, sequence analysis and characterization of cDNA clones coding for
ecerebellin.
A;Reference number: S29328; MUID:93011118; PMID:1396691
A;Accession: S29328
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-246 <PET>
A;Cross-references: EMBL:X66295; NID:g50228; PIDN:CAA46993.1; PID:g50229
C;Superfamily: complement subcomponent C1q chain A; complement C1q carboxyl-terminal
F;122-245/Domain: complement C1q carboxyl-terminal homology <C1Q>

  Query Match             31.1%;  Score 412;  DB 2;  Length 246;
  Best Local Similarity   42.2%;  Pred. No. 5.4e-22;
  Matches  103;  Conservative   33;  Mismatches   84;  Indels   24;  Gaps   10;

Qy          4 LLVLLLLGLAAGSPPLDDNKIPSLC---PGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPG 60
              ||:| || |     ||  ::  : |   || ||:|| ||  |  || |  |  |  |    ||
Db         15 LLLLFLLAL-----PL-RSQASAGCYGIPGMPGMPGAPGKDGHDGLQGPKGEPGIPAVPG 68

Qy         61 APGEKGEGGRPGLPGPRGDPGPRGEA---GPAGPTGPAGECSVPPR------SAFSAKRS 111
                | ||: | ||:|| ||  |||| :   |  || || ||  |  |       | |:  |
Db         69 TQGPKGQKGEPGMPGHRGKNGPRGTSGLPGDPGPRGPPGEPGVEGRYKQKHQSVFTVTRQ 128

Qy        112 ESRVPPPSDAPLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKN 171
              ::  | ::| : |: |: | ||||:  ||||||:|||:||| |: | : |:|   |  |
Db        129 TTQY-PEANALVRFNSVVTNPQGHYNPSTGKFTCEVPGLYYF-VYYTSHTANLCVHLNLN 186

Qy        172 GESIASFFQFFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFL 231
              :|||        |  | ||||::||:  |:||  : | || |:     ::| |||||
Db        187 LARVASFCDHMFN-SKQVS-SGGALLRLQRGDEVW--LSVNDYNGMVGIEGSNSVFSGFL 242

Qy        232 VYSD 235
              :: |
Db        243 LFPD 246


RESULT 6
S13301
collagen alpha 1(X) chain precursor - bovine
C;Species: Bos primigenius taurus (cattle)
C;Date: 21-Nov-1993 #sequence_revision 23-Feb-1996 #text_change 13-Aug-1999
C;Accession: S13301
R;Thomas, J.T.; Kwan, A.P.L.; Grant, M.E.; Boot-Handford, R.P.
Biochem. J. 273, 141-148, 1991
A;Title: Isolation of cDNAs encoding the complete sequence of bovine type X collagen.
A;Reference number: S13301; MUID:91113131; PMID:1703407
A;Accession: S13301
A;Molecule type: mRNA
A;Residues: 1-674 <THO>
A;Cross-references: EMBL:X53556; NID:g263; PIDN:CAA37624.1; PID:g264
C;Genetics:
A;Gene: COL10A1
C;Superfamily: collagen alpha 1(VIII) chain; complement C1q carboxyl-terminal homolog
C;Keywords: coiled coil; extracellular matrix; glycoprotein; homotrimer
F;1-18/Domain: signal sequence #status predicted <SIG>
F;19-674/Product: collagen alpha 1(X) chain #status predicted <MAT>
F;547-673/Domain: complement C1q carboxyl-terminal homology <C1Q>

  Query Match             31.1%;  Score 411.5;  DB 2;  Length 674;
  Best Local Similarity   34.9%;  Pred. No. 1.6e-21;
  Matches   96;  Conservative   34;  Mismatches   84;  Indels   61;  Gaps    6;
```

```
Qy          15 GSPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPGAPGEKGEGGRPGLP 74
               |:| |   |       | |||||  |  |::|:|| :|  |  | || || :|  | ||:|
Db         401 GNPGLPGPKGDPGIAGSPGLPGPVGPAGAKGVPGHNGEAGPRGVPGIPGTRGPIGPPGIP 460

Qy          75 ----------------------------GPRGDPGPRGEA------GPAGPTGPAGECS 99
                                           || | ||||| |      || || || |: :
Db         461 GFPGSKGDVGTPGPPGPAGIAVKGLNGPTGPPGPPGPRGNAGEPGLPGPPGPPGPPGQVA 520

Qy         100 VP--------------------PRSAFSAKRSESRVPPPSDAPLPFDRVLVNEQGHYDA 138
               :|                    | |||:   |::   |    |:|||::| |:| |||
Db         521 LPEDFVKAGQRPFVSANQGVTGMPVSAFTVILSKAY--PAIGTPIPFDKILYNKQQHYDP 578

Qy         139 VTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGESIA-SFFQFFGGWPKPASLSGGAMV 197
               || |||::||:|||: |  |        | |||  :  :: ::  |:   |  || |::
Db         579 RTGIFTCKIPGIYYFSYHIHVKGTHAWVGLYKNGTPVMYTYDEYIKGYLDQA--SGSAVI 636

Qy         198 RLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLV 232
               |   ||||:|:      |:|:     |:||||||
Db         637 DLTENDQVWLQLPNAGSNGLYSPEYVHSSFSGFLV 671


RESULT 7
S31216
collagen alpha 1(X) chain precursor - mouse
C;Species: Mus musculus (house mouse)
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 13-Aug-1999
C;Accession: S31216; S28807; S22215; S30127; I48299; S26397; S31830
R;Kong, R.Y.C.; Kwan, K.M.; Lau, E.T.; Thomas, J.T.; Boot-Handford, R.P.; Grant, M.E.; C
Eur. J. Biochem. 213, 99-111, 1993
A;Title: Intron-exon structure, alternative use of promoter and expression of the mouse
A;Reference number: S31216; MUID:93238750; PMID:8477738
A;Accession: S31216
A;Molecule type: DNA
A;Residues: 1-680 <KON>
A;Cross-references: EMBL:Z21610; NID:g49793; PIDN:CAA79736.1; PID:g49794
R;Elima, K.; Eerola, I.; Rosati, R.; Metsaeranta, M.; Garofalo, S.; Peraelae, M.; de Cro
Biochem. J. 289, 247-253, 1993
A;Title: The mouse collagen X gene: complete nucleotide sequence, exon structure and exp
A;Reference number: S28807; MUID:93143676; PMID:8424763
A;Accession: S28807
A;Molecule type: DNA
A;Residues: 1-285,'A',287-680 <ELI>
A;Cross-references: EMBL:X67348; NID:g50480; PIDN:CAA47763.1; PID:g50481
R;Elima, K.; Metsaeranta, M.; Kallio, J.; Peraelae, M.; Eerola, I.; Garofalo, S.; de Cro
Biochim. Biophys. Acta 1130, 78-80, 1992
A;Title: Specific hybridization probes for mouse alpha-2(IX) and alpha-1(X) collagen mRN
A;Reference number: S22215; MUID:92182017; PMID:1543751
A;Accession: S22215
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 385-450,'K',452-627 <ELA>
A;Cross-references: EMBL:X63013; NID:g49795; PIDN:CAA44741.1; PID:g49796
R;Apte, S.S.; Olsen, B.R.
Matrix 13, 165-179, 1993
A;Title: Characterization of the mouse type X collagen gene.
A;Reference number: S30127; MUID:93261348; PMID:8492743
A;Accession: S30127
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-12,'F',14-26,'S',28-247,'L',249-285,'A',287-305,'F',307-416,'S',418-499,'L
R;Apte, S.S.; Seldin, M.F.; Hayashi, M.; Olsen, B.R.
Eur. J. Biochem. 206, 217-224, 1992
A;Title: Cloning of the human and mouse type X collagen genes and mapping of the mouse t
A;Reference number: I48299; MUID:92267014; PMID:1587271
A;Accession: I48299
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 52-247,'L',249-285,'A',287-305,'F',307-416,'S',418-499,'L',501-566,'C',568,'
A;Cross-references: EMBL:X65121; NID:g50482; PIDN:CAA46237.1; PID:g667031
R;Summers, T.A.; Irwin, M.H.; Mayne, R.; Balian, G.
J. Biol. Chem. 263, 581-587, 1988
A;Title: Monoclonal antibodies to type X collagen. Biosynthetic studies using an antibod
A;Reference number: S26397; MUID:88087150; PMID:2826450
A;Accession: S26397
A;Molecule type: protein
A;Residues: 'SDGYFSQ',24-26,'KQ' <SUM>
C;Genetics:
A;Gene: Col10a-1
A;Map position: 10
A;Introns: 51/3
C;Superfamily: collagen alpha 1(VIII) chain; complement C1q carboxyl-terminal homolog
C;Keywords: coiled coil; extracellular matrix; glycoprotein; homotrimer
F;1-18/Domain: signal sequence #status predicted <SIG>
F;19-680/Product: collagen alpha 1(X) chain #status predicted <MAT>
F;553-679/Domain: complement C1q carboxyl-terminal homology <C1Q>

  Query Match            30.8%;  Score 408;  DB 2;  Length 680;
  Best Local Similarity  33.4%;  Pred. No. 2.9e-21;
  Matches   99;  Conservative   34;  Mismatches   75;  Indels   88;  Gaps    9;

Qy          15 GSPPLDDNKIPSLCPGHPGLP----------------------GTPGHH------GSQ 44
               | | |:  |      |:||||                      | |||:      |
Db         392 GEPGLNGPK------GNPGLPGQKGDPGVGGTPGLRGPVGPVGAKGVPGHNGEAGPRGEP 445

Qy          45 GLPGRDGRDGRDGAPGAPGEKGEGGRPGLPGPR-------------GDPGPRGEA---- 86
               |:||  |  |  | || || ||: | || |||              | |||||  :
Db         446 GIPGTRGPTGPPGVPGFPGSKGDPGNPGAPGPAGIATKGLNGPTGPPGPPGPRGHSGEPG 505

Qy          87 --GPAGPTGPAGECSVP-------------------------PRSAFSAKRSESRVPP 117
                 || || || |:  :|                       | |||:   |::   |
Db         506 LPGPPGPPGPPGQAVMPDGFIKAGQRPRLSGMPLVSANHGVTGMPVSAFTVILSKAY--P 563

Qy         118 PSDAPLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNG-ESIA 176
                 ||:||| :| | | |||  :| |||::||:|||: |  |    :   | |||  ::
Db         564 AVGAPIPFDEILYNRQQHYDPRSGIFTCKIPGIYYFSYHVHVKGTHVWVGLYKNGTPTMY 623

Qy         177 SFFQFFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLV 232
               :: ::  |:   |  || |:: |   ||||:|:   :  |:|:|    |:||||||
Db         624 TYDEYSKGYLDQA--SGSAIMELTENDQVWLQLPNAESNGLYSSEYVHSSFSGFLV 677


RESULT 8
C1HUQC
complement subcomponent C1q chain C precursor - human
N;Alternate names: complement subcomponent C1q gamma chain
C;Species: Homo sapiens (man)
C;Date: 22-May-1981 #sequence_revision 31-May-1996 #text_change 22-May-1998
C;Accession: S14351; A03207
R;Sellar, G.C.; Blake, D.J.; Reid, K.B.M.
Biochem. J. 274, 481-490, 1991
A;Title: Characterization and organization of the genes encoding the A-, B- and C-cha
A;Reference number: S14350; MUID:91174759; PMID:1706597
A;Accession: S14351
A;Status: not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 1-245 <SEL>
R;Reid, K.B.M.
Biochem. J. 179, 367-371, 1979
A;Title: Complete amino acid sequences of the three collagen-like regions present in
A;Reference number: A90304; MUID:80020137; PMID:486087
A;Accession: A03207
A;Molecule type: protein
A;Residues: 29-56,'P',58-65,'K',67-71,'P',73-83,'K',85-86,'D',88-89,'N',91-122 <REI>
C;Comment: The first component of complement is a calcium-dependent complex of the th
ivation of C1r (enzyme), C1s (proenzyme), and the other eight components of complemen
C;Comment: The C1q subcomponent is composed of nine subunits, six of which are disulf
 dimers of the C chain. Equimolar amounts of the A, B, and C chains are found after r
C;Genetics:
A;Gene: GDB:C1QG
A;Cross-references: GDB:128132; OMIM:120575
A;Map position: 1p36.3-1p34.1
A;Introns: 60/3
C;Superfamily: complement subcomponent C1q chain A; complement C1q carboxyl-terminal
C;Keywords: complement pathway; glycoprotein; homodimer; hydroxylysine; hydroxyprolin
F;1-28/Domain: signal sequence #status predicted <SIG>
```

```
F;29-245/Product: complement subcomponent C1q chain B #status predicted <MAT>
F;31-114/Domain: collagenous, triple helix <COL>
F;121-244/Domain: complement C1q carboxyl-terminal homology <C1Q>
F;32/Disulfide bonds: interchain #status experimental
F;36,39,42,45,54,63,81,93,96,99,105/Modified site: 4-hydroxyproline (Pro) #status experi
F;57,72,75/Modified site: 5-hydroxylysine (Lys) #status experimental
F;75/Binding site: carbohydrate (Lys) (covalent) #status experimental

  Query Match             30.1%;  Score 399;  DB 1;  Length 245;
  Best Local Similarity   39.8%;  Pred. No. 4.3e-21;
  Matches  100;  Conservative   33;  Mismatches   84;  Indels   34;  Gaps   10;

Qy          1 MRPLLVLLLLGLAAGSPPLDDNKIPSLC---PGHPGLPGTPGHHGSQGLPGRDGRDGRDG 57
              :: ||:|||| |          :  : |   || ||||| ||  |  ||||  |  |
Db         13 LKLLLLLLLLALR--------GQANTGCYGIPGMPGLPGAPGKDGYDGLPGPKGEPGIPA 64

Qy         58 APGAPGEKGEGGRPGLPGPRGDPGPRGEAGPAGPTGPAGECSVPP---------RSAFSA 108
               ||  | ||: | |||||  |  || |  |  | .|| |     |          :| |:
Db         65 IPGIRGPKGQKGEPGLPGHPGKNGPMGPPGMPGVPGPMGIPGEPGEEGRYKQKFQSVFTV 124

Qy        109 KRSESRVPPPSDAPLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDL 168
              | ::  ||  :: : |: || | || ||  ||||||:|||:|||  ||: : |:|   |
Db        125 TR-QTHQPPAPNSLIRFNAVLTNPQGDYDTSTGKFTCKVPGLYYFVYHAS-HTANLCVLL 182

Qy        169 VKNGESIASFFQFFGGWPKPASL-SGGAMVRLEPEDQVWVQVGVGDY---IGIYASIKTD 224
              ::|  : :   | |   |   : ||| ::||:  ::||  : | ||   :||  |   |
Db        183 YRSGVKVVT---FCGHTSKTNQVNSGGVLLRLQVGEEVW--LAVNDYYDMVGIQGS---D 234

Qy        225 STFSGFLVYSD 235
              | ||||:: |
Db        235 SVFSGFLLFPD 245


RESULT 9
A34246
collagen alpha 1(VIII) chain precursor - rabbit
C;Species: Oryctolagus cuniculus (domestic rabbit)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C;Accession: A34246
R;Yamaguchi, N.; Benya, P.D.; van der Rest, M.; Ninomiya, Y.
J. Biol. Chem. 264, 16022-16029, 1989
A;Title: The cloning and sequencing of alphal(VIII) collagen cDNAs demonstrate that type
omains similar to those of type X collagen.
A;Reference number: A34246; MUID:89380199; PMID:2476437
A;Accession: A34246
A;Molecule type: mRNA
A;Residues: 1-744 <YAM>
A;Cross-references: GB:J05042; NID:g164895; PIDN:AAA31204.1; PID:g164896
C;Superfamily: collagen alpha 1(VIII) chain; complement C1q carboxyl-terminal homology
F;1-20/Domain: signal sequence #status predicted <SIG>
F;21-744/Product: collagen alpha 1(VIII) chain #status predicted <MAT>
F;21-117/Region: amino-terminal nonhelical
F;118-571/Region: interrupted helical
F;572-744/Region: carboxyl-terminal nonhelical
F;617-743/Domain: complement C1q carboxyl-terminal homology <C1Q>

  Query Match             28.2%;  Score 374;  DB 1;  Length 744;
  Best Local Similarity   34.4%;  Pred. No. 7.6e-19;
  Matches   93;  Conservative   36;  Mismatches   79;  Indels   62;  Gaps    8;

Qy         24 IPSLC--PGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPGAPGEKGEGGRPGLP------- 74
              :| |    | ||:||  |  |  |:||  |  |  | || || ||| | || |
Db        475 VPGLLGPKGEPGIPGDQGLQGPPGIPGITGPSGPIGPPGIPGPKGEPGLPGPPGFPGVGK 534

Qy         75 -------GPRGDP---GPRGE---AGPAGPTGPAGECSVPPRS----------------- 104
                     || | |   ||:|:    || || || |  :| | :
Db        535 PGVAGLHGPPGKPGALGPQGQPGLPGPPGPPGPPGPPAVMPPTPAPQGEYLPDMGLGIDG 594

Qy        105 -----AFSAKRSESRVP--------------PPSDAPLPFDRVLVNEQGHYDAVTGKFT 144
                  |::||: ::  |               ||  ||: |||:| | : :|:  || ||
Db        595 VKTPHAYAAKKGKNGGPAYEMPAFTAELTAPFPPVGAPIKFDRLLYNGRQNYNPQTGIFT 654

Qy        145 CQVPGVYYFAVHATVYRASLQFDLVKNGESIA-SFFQFFGGWPKPASLSGGAMVRLEPED 203
              |:|||||||| |       ::   | || | :  :: ::  |:   |  || |:: | | |
Db        655 CEVPGVYYFAYHVHCKGGNVWVALFKNNEPVMYTYDEYKKGFLDQA--SGSAVLLLRPGD 712

Qy        204 QVWVQVGVGDYIGIYASIKTDSTFSGFLVY 233
              :|::|:      |:||     |:|||:|:|
Db        713 RVFLQMPSEQAAGLYAGQYVHSSFSGYLLY 742


RESULT 10
S23298
collagen alpha 1(VIII) chain - chicken
C;Species: Gallus gallus (chicken)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C;Accession: S23298
R;Ninomiya, Y.; Castagnola, P.; Gerecke, D.; Gordon, M.K.; Jacenko, O.; LuValle, P.;
maguchi, N.; Olsen, B.R.
in Extracellular Matrix Genes, Sandell L.J. and Boyd C.D., eds., pp.79-114, Academic
A;Title: The molecular biology of collagens with short triple-helical domains.
A;Reference number: S22243
A;Accession: S23298
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-744 <NIN>
C;Superfamily: collagen alpha 1(VIII) chain; complement C1q carboxyl-terminal homolog
F;617-743/Domain: complement C1q carboxyl-terminal homology <C1Q>

  Query Match             27.9%;  Score 370;  DB 1;  Length 744;
  Best Local Similarity   34.4%;  Pred. No. 1.4e-18;
  Matches   93;  Conservative   35;  Mismatches   80;  Indels   62;  Gaps    8;

Qy         24 IPSLC--PGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPGAPGEKGEGGRPGLP------- 74
              :| |    | ||:||  |  |  |:||  |  |  | || || ||| | || |
Db        475 VPGLLGPKGEPGIPGDQGLQGPPGIPGITGPSGPIGPPGIPGPKGEPGLPGPPGFPGVGK 534

Qy         75 -------GPRGDP---GPRGE---AGPAGPTGPAGECSVPPRS----------------- 104
                     || | |   ||:|:    || || || |  :| | :
Db        535 PGVAGLHGPPGKPGALGPQGQPGLPGPPGPPGPPGPPAVMPPTPAPQGEYLPDMGLGIDG 594

Qy        105 -----AFSAKRSESRVP--------------PPSDAPLPFDRVLVNEQGHYDAVTGKFT 144
                  |::||: ::  |               ||  ||: |||:| | : :|:  || ||
Db        595 VKTPHAYAAKKGKNGGPAYEMPAFTAELTAPFPPVGAPIKFDRLLYNGRQNYNPQTGIFT 654

Qy        145 CQVPGVYYFAVHATVYRASLQFDLVKNGESIA-SFFQFFGGWPKPASLSGGAMVRLEPED 203
              |:|||||||| |       ::   | || | :  :: ::  |:   |  || |:: | | |
Db        655 CEVPGVYYFAYHVHCKGGNVWVALFKNNEPVMYTYDEYKKGFLDQA--SGSAVLLLRPGD 712

Qy        204 QVWVQVGVGDYIGIYASIKTDSTFSGFLVY 233
              :|::|      |:||     |:|||:|:|
Db        713 RVFLQNPSEQAAGLYAGQYVHSSFSGYLLY 742


RESULT 11
S23779
collagen alpha 1(VIII) chain - mouse
C;Species: Mus musculus (house mouse)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 16-Jun-2000
C;Accession: S23779
R;Muragaki, Y.; Shiota, C.; Inoue, M.; Ooshima, A.; Olsen, B.R.; Ninomiya, Y.
Eur. J. Biochem. 207, 895-902, 1992
A;Title: Alpha-1(VIII)-collagen gene transcripts encode a short-chain collagen polype
A;Reference number: S23779; MUID:92362626; PMID:1499564
A;Accession: S23779
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-743 <MUR>
A;Cross-references: EMBL:X66976; NID:g50493; PIDN:CAA47387.1; PID:g1359953
C;Superfamily: collagen alpha 1(VIII) chain; complement C1q carboxyl-terminal homolog
F;616-742/Domain: complement C1q carboxyl-terminal homology <C1Q>

  Query Match             27.4%;  Score 363.5;  DB 1;  Length 743;
  Best Local Similarity   32.8%;  Pred. No. 4.1e-18;
  Matches   90;  Conservative   35;  Mismatches   78;  Indels   71;  Gaps    7;
```

```
Qy          29 PGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPGAPGEKGEGGRPGLPGPRGDP-------- 80
               || ||:||  |  |  |:|| |  |  | || |  |  | ||:|||:|:|
Db         470 PGLPGVPGLLGPKGEPGIPGDQGLQGPPGIPGIVGPSGPIGPPGIPGPKGEPGLPGPPGF 529

Qy          81 --------------------GPRGE---AGPAGPTGPAGECSVPPR------------ 103
                                   ||:|:    || || || |  :| |
Db         530 PGVGKPGVAGLHGPPGKPGALGPQGQPGLPGPPGPPGPPGPPAVMPTPSPQGEYLPDMGL 589

Qy         104 --------SAFSAKR--------------SESRVP-PPSDAPLPFDRVLVNEQGHYDAVT 140
                     |:: |:                 :|  || ||  ||: ||::| | : :|:  |
Db         590 GIDGVKTPHAYAGKKGKHGGPAYEMPAFTAELTVPFPPVGAPVKFDKLLYNGRQNYNPQT 649

Qy         141 GKFTCQVPGVYYFAVHATVYRASLQFDLVKNGESIA-SFFQFFGGWPKPASLSGGAMVRL 199
               | |||:|||||||| |      ::   | || | :  :: ::  |:   |  || |:: |
Db         650 GIFTCEVPGVYYFAYHVHCKGGNVWVALFKNNEPMMYTYDEYKKGFLDQA--SGSAVLLL 707

Qy         200 EPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVY 233
                | |||::|       |:||     |:|||:|:|
Db         708 RPGDQVFLQNPFEQAAGLYAGQYVHSSFSGYLLY 741


RESULT 12
A57131
collagen alpha 2(VIII) chain - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 03-Nov-1995 #sequence_revision 03-Nov-1995 #text_change 13-Aug-1999
C;Accession: A57131
R;Muragaki, Y.; Jacenko, O.; Apte, S.; Mattei, M.G.; Ninomiya, Y.; Olsen, B.R.
J. Biol. Chem. 266, 7721-7727, 1991
A;Title: The alpha2(VIII) collagen gene. A novel member of the short chain collagen fami
A;Reference number: A57131; MUID:91210292; PMID:2019595
A;Accession: A57131
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-635 <MUR>
A;Cross-references: GB:M60832; NID:g177178; PIDN:AAA62822.1; PID:g177179
C;Genetics:
A;Gene: GDB:COL8A2
A;Cross-references: GDB:127812; OMIM:120252
A;Map position: 1p34.3-1p32.3
C;Superfamily: collagen alpha 1(VIII) chain; complement C1q carboxyl-terminal homology
F;1-11/Domain: amino-terminal nonhelical (fragment) #status predicted <NC2>
F;12-468/Region: interrupted helical
F;469-635/Domain: carboxyl-terminal nonhelical #status predicted <NC7>
F;508-634/Domain: complement C1q carboxyl-terminal homology <C1Q>

  Query Match             27.4%;  Score 363;  DB 2;  Length 635;
  Best Local Similarity   38.6%;  Pred. No. 3.8e-18;
  Matches   93;  Conservative   25;  Mismatches   95;  Indels   28;  Gaps    7;

Qy          11 GLAAGSPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPGAPGEKGEGGR 70
               ||   : |:    :| |  | ||||| ||    :|  |  |  | || ||  |  |
Db         401 GLQGPAGPIGPQGLPGL-KGEPGLPGPPG----EGRAGEPGTAGPRGPPGVPGSPGITGP 455

Qy          71 PGLPGPRGDPGPRGEAGPAG---PTGPA--------------GECSVPPRSAFSAKRSE 112
               |||||| | ||   | | ||   | |                || |     ||:|  :
Db         456 PGLPGPPGAPGAFDETGIAGLHLPNGGVEGAVLGKGGKPQFGLGELSAHATPAFTAVLTS 515

Qy         113 SRVPPPSDAPLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNG 172
                  | |  |: ||| | |    |:  || ||| | |||||| |  |   ::   | ||
Db         516 PL--PASGMPVKFDRTLYNGHSGYNPATGIFTCPVGGVYYFAYHVHVKGTNVWVALYKNN 573

Qy         173 -ESIASFFQFFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFL 231
                 :  :: ::  |:   |  ||||:::| | ||||||:      |:|::    |:|||||
Db         574 VPATYTYDEYKKGYLDQA--SGGAVLQLRPNDQVWVQIPSDQANGLYSTEYIHSSFSGFL 631

Qy         232 V 232
               :
Db         632 L 632


RESULT 13
C1HUQB
```

```
complement subcomponent C1q chain B precursor [validated] - human
N;Alternate names: complement subcomponent C1q beta chain
C;Species: Homo sapiens (man)
C;Date: 22-May-1981 #sequence_revision 31-May-1996 #text_change 08-Dec-2000
C;Accession: B23422; A23422; B90304; A90301; B90315; A03206
R;Reid, K.B.M.
Biochem. J. 231, 729-735, 1985
A;Title: Molecular cloning and characterization of the complementary DNA and gene cod
A;Reference number: A23422; MUID:86076906; PMID:3000358
A;Accession: B23422
A;Molecule type: DNA
A;Residues: 'HS',1-32 <REI>
A;Note: the authors translated the codon AGT for the second position as Arg; they wer
A;Accession: A23422
A;Molecule type: mRNA
A;Residues: 28-253 <RE1>
A;Cross-references: EMBL:X03084
A;Note: the authors translated the codon ACA for residue 46 as Ile
R;Reid, K.B.M.
Biochem. J. 179, 367-371, 1979
A;Title: Complete amino acid sequences of the three collagen-like regions present in
A;Reference number: A90304; MUID:80020137; PMID:486087
A;Accession: B90304
A;Molecule type: protein
A;Residues: 'E',29-84,'D',86-99,'P',101-135 <RE5>
R;Reid, K.B.M.; Thompson, E.O.P.
Biochem. J. 173, 863-868, 1978
A;Title: Amino acid sequence of the N-terminal 108 amino acid residues of the B chain
A;Reference number: A90301; MUID:79041552; PMID:708376
A;Accession: A90301
A;Molecule type: protein
A;Residues: 28-99,'P',101-195 <RE3>
R;Reid, K.B.M.; Gagnon, J.; Frampton, J.
Biochem. J. 203, 559-569, 1982
A;Title: Completion of the amino acid sequences of the A and B chains of subcomponent
A;Reference number: A90315; MUID:82283890; PMID:6981411
A;Accession: B90315
A;Molecule type: protein
A;Residues: 136-253 <RE4>
A;Note: 176-Glx may also be present
C;Comment: The first component of complement is a calcium-dependent complex of the th
ivation of C1r (enzyme), C1s (proenzyme), and the other eight components of complemen
C;Comment: The C1q subcomponent is composed of nine subunits, six of which are disulf
(see PIR:C1HUQC) chain. Equimolar amounts of the A, B, and C chains are found after r
C;Genetics:
A;Gene: GDB:C1QB
A;Cross-references: GDB:119043; OMIM:120570
A;Map position: 1p36.3-1p34.1
C;Superfamily: complement subcomponent C1q chain A; complement C1q carboxyl-terminal
C;Keywords: complement pathway; glycoprotein; heterodimer; hydroxylysine; hydroxyprol
F;1-27/Domain: signal sequence #status predicted <SIG>
F;28-253/Product: complement subcomponent C1q chain B #status experimental <MAT>
F;33-116/Domain: collagenous, triple helix <COL>
F;123-249/Domain: complement C1q carboxyl-terminal homology <C1Q>
F;28/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status experi
F;31/Disulfide bonds: interchain (to chain A-26) #status experimental
F;35,38,41,53,56,65,83,86,101,104,107/Modified site: 4-hydroxyproline (Pro) #status e
F;59,62,77,92,98,110/Modified site: 5-hydroxylysine (Lys) #status experimental
F;59,62,98,110/Binding site: carbohydrate (Lys) (covalent) #status experimental

  Query Match             27.3%;  Score 362;  DB 1;  Length 253;
  Best Local Similarity   36.9%;  Pred. No. 1.7e-18;
  Matches   90;  Conservative   36;  Mismatches  100;  Indels   18;  Gaps    6;

Qy           4 LLVLLLLGLAAGSPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPGAPG 63
               |:|||||||   |        |   || ||:|||||  |  | ||  |  |  |  |  |
Db          13 LMLLLLLGLIDISQAQLSCTGPPAIPGIPGIPGTPGPDGQPGTPGIKGEKGLPGLAGDHG 72

Qy          64 EKGEGGRPGLPGPRGDPGPRGEAGP---------AGPTGPAGECSVPPRSAFSAKRSESR 114
               | || | ||:||  |  ||:|  ||          || | :|:     : |||| |:
Db          73 EFGEKGDPGIPGNPGKVGPKGPMGPKGGPGAPGAPGPKGESGDYKATQKIAFSATRT-IN 131

Qy         115 VPPPSDAPLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGE- 173
```

```
           ||   |  : || |: |   :|:  :|||||:|||:|||  ||:  | :|  :|::  |
Db        132 VPLRRDQTIRFDHVITNMNNNYEPRSGKFTCKVPGLYYFTYHAS-SRGNLCVNLMRGRER 190

Qy        174 --SIASFFQFFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFL 231
              : :|  :  :      :|| :::||  : |::|    |  :      :| |||||
Db        191 AQKVVTFCDY--AYNTFQVTTGGMVLKLEQGENVFLQ--ATDKNSLLGMEGANSIFSGFL 246

Qy        232 VYSD 235
              :: |
Db        247 LFPD 250


RESULT 14
S49158
complement protein C1q beta chain precursor - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 16-Feb-1995 #sequence_revision 12-May-1995 #text_change 20-Aug-1999
C;Accession: S49158
R;Schwaeble, W.; Petry, F.; Loos, M.
submitted to the EMBL Data Library, March 1993
A;Description: cDNA cloning and expression of the mRNA encoding the B-chain of rat C1q.
A;Reference number: S49158
A;Accession: S49158
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-253 <SCH>
A;Cross-references: EMBL:X71127; NID:g510191; PIDN:CAA50440.1; PID:g510192
C;Superfamily: complement subcomponent C1q chain A; complement C1q carboxyl-terminal hom
F;121-249/Domain: complement C1q carboxyl-terminal homology <C1Q>

  Query Match             27.2%;  Score 360.5;  DB 2;  Length 253;
  Best Local Similarity   36.1%;  Pred. No. 2.2e-18;
  Matches   95;  Conservative   43;  Mismatches   76;  Indels   49;  Gaps   13;

Qy          1 MRPLLVLLLLGLAAGSPPLDDNKIPSLC---PGHPGLPGTPGHHGSQGLPGRDGRDGRDG 57
              : ||| ||||||      |  :   | |   || ||:|| ||  || | ||  |  |  |
Db          9 LTPLL-LLLLGL------LHVSWAQSSCTGSPGIPGVPGIPGVPGSDGKPGTPGIKGEKG 61

Qy         58 APGAP------GEKGEGGRPGLP---GPRGDPGPRGEAGPAGPTGP---AGECSVPPRSA 105
               ||        ||||: | ||:|   ||:|  ||:|  || || ||   :|:     : |
Db         62 LPGLAGDHGELGEKGDAGIPGIPGKVGPKGPVGPKGAPGPPGPRGPKGGSGDYKATQKVA 121

Qy        106 FSAKRSESRVPPPSDAPLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQ 165
              ||| |: :    |: | : |::|: |   :|:  :|||||:|||:|||  ||:  | :|
Db        122 FSALRTVNSALRPNQA-IRFEKVITNVNDNYEPRSGKFTCKVPGLYYFTYHAS-SRGNLC 179

Qy        166 FDLVKNG-----ESIASF-------FQFFGGWPKPASLSGGAMVRLEPEDQVWVQ-VGVG 212
              ::|:       : : :|       ||           :|| :::|| |: | :|
Db        180 VNIVRGRDRDRMQKVLTFCDYAQNTFQV---------TTGGVVLKLEQEEVVHLQATDKN 230

Qy        213 DYIGIYASIKTDSTFSGFLVYSD 235
                :|:  :   :| |:|||:: |
Db        231 SLLGVEGA---NSIFTGFLLFPD 250


RESULT 15
S15435
collagen alpha 1(VIII) chain precursor - human
C;Species: Homo sapiens (man)
C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 13-Aug-1999
C;Accession: S15435
R;Muragaki, Y.; Mattei, M.G.; Yamaguchi, N.; Olsen, B.R.; Ninomiya, Y.
Eur. J. Biochem. 197, 615-622, 1991
A;Title: The complete primary structure of the human alpha-1(VIII) chain and assignment
A;Reference number: S15435; MUID:91231001; PMID:2029894
A;Accession: S15435
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-744 <MUR>
A;Cross-references: EMBL:X57527; NID:g30081; PIDN:CAA40748.1; PID:g30082
C;Genetics:
A;Gene: GDB:COL8A1
A;Cross-references: GDB:128104; OMIM:120251
A;Map position: 3q11.1-3q13.2
C;Superfamily: collagen alpha 1(VIII) chain; complement C1q carboxyl-terminal homolog
F;1-20/Domain: signal sequence #status predicted <SIG>
F;21-744/Product: collagen alpha 1(VIII) chain #status predicted <MAT>
F;21-117/Region: amino-terminal nonhelical
F;118-571/Region: interrupted helical
F;572-744/Region: carboxyl-terminal nonhelical
F;617-743/Domain: complement C1q carboxyl-terminal homology <C1Q>

  Query Match             27.1%;  Score 359;  DB 2;  Length 744;
  Best Local Similarity   33.0%;  Pred. No. 8.5e-18;
  Matches   89;  Conservative   33;  Mismatches   86;  Indels   62;  Gaps    6;

Qy         24 IPSLC--PGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPGAPGEKGEGGRPGLP------- 74
              :| |    | ||:||  |  |  |:||  |  |  | || || ||| | || |
Db        475 VPGLLGPKGEPGIPGDQGLQGPPGIPGIGGPSGPIGPPGIPGPKGEPGLPGPPGFPGIGK 534

Qy         75 ----------------GPRGDPGPRGEAGPAGPTGP------------------AGECS 99
                             ||:| ||  |  || || ||                    |
Db        535 PGVAGLHGPPGKPGALGPQGQPGLPGPPGPPGPPGPPAVMPPTPPPQGEYLPDMGLGIDG 594

Qy        100 VPPRSAFSAKRSESRVP--------------PPSDAPLPFDRVLVNEQGHYDAVTGKFT 144
              | |  |  ||: ::  |               ||   |: |:::| | : :|:  || ||
Db        595 VKPPHATGAKKGKNGGPAYEMPAFTAELTAPFPPVGGPVKFNKLLYNGRQNYNPQTGIFT 654

Qy        145 CQVPGVYYFAVHATVYRASLQFDLVKNGESIA-SFFQFFGGWPKPASLSGGAMVRLEPED 203
              |:|||||||| |       ::   | || | :  :: ::  |:   |  || |:: | | |
Db        655 CEVPGVYYFAYHVHCKGGNVWVALFKNNEPVMYTYDEYKKGFLDQA--SGSAVLLLRPGD 712

Qy        204 QVWVQVGVGDYIGIYASIKTDSTFSGFLVY 233
              :|::|:      |:||     |:|||:|:|
Db        713 RVFLQMPSEQAAGLYAGQYVHSSFSGYLLY 742


Search completed: June 18, 2003, 14:53:59
Job time : 53 secs
```

GenCore version 5.1.6


OM protein - protein search, using sw model

Run on: June 18, 2003, 14:40:55 ; Search time 83 Seconds
(without alignments)
390.119 Million cell updates/sec

Title: US-09-943-851A-42
Perfect score: 1325
Sequence: 1 MRPLLVLLLLGLAAGSPPLD.....DSTFSGFLVYSDWHSSPVFA 243

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 1325 | 100.0 | 243 | 20 | AAY06481 | Human tumour-assoc |
| 2 | 1325 | 100.0 | 243 | 20 | AAY17827 | Human PRO344 prote |
| 3 | 1325 | 100.0 | 243 | 20 | AAW97984 | Human adipocyte-sp |
| 4 | 1325 | 100.0 | 243 | 21 | AAB33461 | Human PRO344 prote |
| 5 | 1325 | 100.0 | 243 | 21 | AAY71468 | Human PRO344 prote |
| 6 | 1325 | 100.0 | 243 | 21 | AAY93688 | Amino acid sequenc |
| 7 | 1325 | 100.0 | 243 | 21 | AAB01318 | Human PRO344 polyp |
| 8 | 1325 | 100.0 | 243 | 22 | AAU12352 | Human PRO344 polyp |
| 9 | 1325 | 100.0 | 243 | 22 | AAB65815 | Human TANGO 253 SE |
| 10 | 1325 | 100.0 | 243 | 22 | AAB49593 | Human adipocyte co |
| 11 | 1325 | 100.0 | 243 | 22 | AAB49599 | Human adipocyte co |
| 12 | 1323 | 99.8 | 243 | 22 | AAB65891 | Human secreted pro |
| 13 | 1321 | 99.7 | 243 | 22 | AAB65888 | Human secreted pro |
| 14 | 1321 | 99.7 | 243 | 22 | AAB65889 | Human secreted pro |
| 15 | 1321 | 99.7 | 243 | 22 | AAB65890 | Human secreted pro |
| 16 | 1271 | 95.9 | 243 | 21 | AAY76040 | Rat skin cell prot |
| 17 | 1271 | 95.9 | 243 | 22 | AAB55979 | Skin cell protein, |
| 18 | 1271 | 95.9 | 243 | 23 | ABB72179 | Rat protein isolat |
| 19 | 1258 | 94.9 | 243 | 22 | AAB65820 | Murine TANGO 253 S |
| 20 | 1256 | 94.8 | 228 | 22 | AAB65816 | Human mature TANGO |
| 21 | 1256 | 94.8 | 243 | 22 | AAB65899 | Murine secreted pr |
| 22 | 1254 | 94.6 | 243 | 22 | AAB65897 | Murine secreted pr |
| 23 | 1254 | 94.6 | 243 | 22 | AAB65898 | Murine secreted pr |
| 24 | 1243.5 | 93.8 | 242 | 22 | AAB65896 | Murine secreted pr |
| 25 | 1200 | 90.6 | 228 | 22 | AAB65821 | Murine mature TANG |
| 26 | 1183 | 89.3 | 220 | 22 | ABG12724 | Novel human diagno |
| 27 | 1048 | 79.1 | 201 | 23 | AAO21663 | Human secreted pro |
| 28 | 675 | 50.9 | 128 | 22 | AAB65819 | Human TANGO 253 C1 |
| 29 | 650 | 49.1 | 128 | 22 | AAB65824 | Murine TANGO 253 C |
| 30 | 478 | 36.1 | 151 | 22 | ABG12723 | Novel human diagno |
| 31 | 449.5 | 33.9 | 225 | 23 | ABB80583 | Human sbg1033026C1 |
| 32 | 442.5 | 33.4 | 333 | 23 | ABB80582 | Human sbg1033026C1 |
| 33 | 433.5 | 32.7 | 247 | 18 | AAW09107 | Murine adipocyte c |
| 34 | 433.5 | 32.7 | 247 | 22 | AAE05528 | Mouse OBG3 protein |
| 35 | 433.5 | 32.7 | 247 | 23 | ABB08222 | Mouse acrp30 prote |
| 36 | 425 | 32.1 | 244 | 18 | AAW09108 | Human adipocyte co |
| 37 | 424.5 | 32.0 | 247 | 22 | AAE05527 | Mouse OBG3 protein |
| 38 | 424.5 | 32.0 | 247 | 23 | ABB08221 | Mouse adipoQ prote |
| 39 | 424 | 32.0 | 244 | 20 | AAY21807 | Adipose most abund |
| 40 | 424 | 32.0 | 244 | 21 | AAB30233 | Human adipocyte co |
| 41 | 424 | 32.0 | 244 | 21 | AAY71035 | Human APM1 (Adipos |
| 42 | 424 | 32.0 | 244 | 22 | AAE05529 | Human OBG3 protein |
| 43 | 424 | 32.0 | 244 | 22 | AAB65828 | Human adipocyte co |
| 44 | 424 | 32.0 | 244 | 22 | AAB49592 | Human ACRP30 prote |
| 45 | 424 | 32.0 | 244 | 22 | AAB49598 | Human ACRP30 prote |

ALIGNMENTS

RESULT 1
AAY06481
ID AAY06481 standard; Protein; 243 AA.
XX
AC AAY06481;
XX
DT 27-SEP-1999 (first entry)
XX
DE Human tumour-associated protein PRO344.
XX
KW PRO344; UNQ303; cancer; tumour; diagnosis; therapy; human.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Peptide 1..15
FT /note= "signal peptide"
FT Protein 16..243
FT /note= "mature protein"
FT Modified-site 68..215
FT /note= "N-myristoylated"
FT Modified-site 216..243
FT /note= "N-myristoylated"
XX
PN WO9935170-A2.
XX
PD 15-JUL-1999.
XX
PF 05-JAN-1999; 99WO-US00106.
XX
PR 20-NOV-1998; 98US-0109304.
PR 05-JAN-1998; 98US-0070440.
PR 29-APR-1998; 98US-0083500.

```
PR   22-MAY-1998;   98US-0086414.
PR   10-JUN-1998;   98US-0088742.
PR   10-NOV-1998;   98US-0107783.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Botstein D,  Goddard A,  Gurney AL,  Hillan KJ,  Lawrence DA;
PI   Roy MA,  Wood WI;
XX
DR   WPI; 1999-430385/36.
DR   N-PSDB; AAX87258.
XX
PT   Antibody against proteins expressed in neoplastic cells, useful for
PT   tumor diagnosis and treatment
XX
PS   Example 1; Fig 10; 162pp; English.
XX
CC   This sequence represents human PRO344 (UNQ303), a protein encoded
CC   by the novel cDNA clone DNA40592 (see AAX87258).  Amplification of
CC   DNA40592 was observed in primary lung tumours and in primary colon
CC   tumours, suggesting a significant role in tumour formation and
CC   growth.  Antagonists (e.g. antibodies) directed to PRO344 may have
CC   use in cancer therapy.  The invention identifies 14 genes (see
CC   AAX87254-67) that are amplified in the genome of tumour cells.  Such
CC   amplification is expected to be associated with overexpression of
CC   the gene product and to contribute to tumorigenesis.  The encoded
CC   proteins (see AAY06477-90) may be useful targets for the diagnosis
CC   and/or treatment (including prevention) of certain cancers, and may
CC   act as predictors of the prognosis of tumour treatment.  Antibodies
CC   that bind the proteins are claimed and used in claimed cancer
CC   diagnostic kits.
XX
SQ   Sequence   243 AA;

  Query Match             100.0%;  Score 1325;  DB 20;  Length 243;
  Best Local Similarity   100.0%;  Pred. No. 5.3e-103;
  Matches  243;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MRPLLVLLLLGLAAGSPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MRPLLVLLLLGLAAGSPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPG 60

Qy         61 APGEKGEGGRPGLPGPRGDPGPRGEAGPAGPTGPAGECSVPPRSAFSAKRSESRVPPPSD 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 APGEKGEGGRPGLPGPRGDPGPRGEAGPAGPTGPAGECSVPPRSAFSAKRSESRVPPPSD 120

Qy        121 APLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGESIASFFQ 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 APLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGESIASFFQ 180

Qy        181 FFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWHSSP 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 FFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWHSSP 240

Qy        241 VFA 243
              |||
Db        241 VFA 243


RESULT 2
AAY17827
ID   AAY17827 standard; Protein; 243 AA.
XX
AC   AAY17827;
XX
DT   12-AUG-1999  (first entry)
XX
DE   Human PRO344 protein sequence.
XX
KW   Human; PRO protein; tumour necrosis factor family; TNF; cytokine;
KW   secreted protein; transmembrane protein; inflammation disorder.
XX
OS   Homo sapiens.
XX
PN   WO9928462-A2.
XX
PD   10-JUN-1999.
XX
PF   01-DEC-1998;   98WO-US25108.
XX
PR   25-FEB-1998;   98US-0075945.
PR   03-DEC-1997;   97US-0067411.
PR   11-DEC-1997;   97US-0069278.
PR   11-DEC-1997;   97US-0069334.
PR   11-DEC-1997;   97US-0069335.
PR   12-DEC-1997;   97US-0069425.
PR   16-DEC-1997;   97US-0069694.
PR   16-DEC-1997;   97US-0069696.
PR   16-DEC-1997;   97US-0069702.
PR   17-DEC-1997;   97US-0069870.
PR   17-DEC-1997;   97US-0069873.
PR   18-DEC-1997;   97US-0068017.
PR   05-JAN-1998;   98US-0070440.
PR   09-FEB-1998;   98US-0074086.
PR   09-FEB-1998;   98US-0074092.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Baker KP,  Chen J,  Goddard A,  Gurney AL,  Wood WI;
PI   Yuan J;
XX
DR   WPI; 1999-371118/31.
DR   N-PSDB; AAX80052.
XX
PT   Nucleic acids encoding PRO secreted and transmembrane proteins
XX
PS   Claim 12; Fig 21; 123pp; English.
XX
CC   The present invention describes nucleic acids encoding PRO secreted and
CC   transmembrane proteins used therapeutically. The PRO proteins have
CC   cytostatic, anti-inflammatory, anti-proliferative and immunosuppressive
CC   activity. The proteins and polynucleotides can be used in therapy,
CC   identification of homologues, raising antibodies and design of probes
CC   and primers. They can be used in a range of diseases related to proteins
CC   that they have homology with, e.g. a PRO protein having homology to
CC   complement proteins may be used in inflammatory responses.
XX
SQ   Sequence   243 AA;

  Query Match             100.0%;  Score 1325;  DB 20;  Length 243;
  Best Local Similarity   100.0%;  Pred. No. 5.3e-103;
  Matches  243;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MRPLLVLLLLGLAAGSPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MRPLLVLLLLGLAAGSPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPG 60

Qy         61 APGEKGEGGRPGLPGPRGDPGPRGEAGPAGPTGPAGECSVPPRSAFSAKRSESRVPPPSD 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 APGEKGEGGRPGLPGPRGDPGPRGEAGPAGPTGPAGECSVPPRSAFSAKRSESRVPPPSD 120

Qy        121 APLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGESIASFFQ 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 APLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGESIASFFQ 180

Qy        181 FFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWHSSP 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 FFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWHSSP 240

Qy        241 VFA 243
              |||
Db        241 VFA 243


RESULT 3
AAW97984
```

```
ID   AAW97984 standard; Protein; 243 AA.
XX
AC   AAW97984;
XX
DT   21-JUN-1999  (first entry)
XX
DE   Human adipocyte-specific protein zsig39.
XX
KW   Adipocyte-specific protein; zsig39; human; fatty acid metabolism;
KW   energy balance; nutrition; antimicrobial.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1..15
FT                   /note= "signal peptide, alternatively the signal
FT                          peptide comprises residues 1..18"
FT   Protein         16..243
FT                   /note= "mature protein, alternatively the mature
FT                          protein comprises residues 19..243
FT                          (specifically claimed in Claim 4)"
FT   Domain          30..96
FT                   /note= "collagen-like domain"
FT   Domain          98..243
FT                   /note= "globular domain"
FT   Region          105..109
FT                   /note= "beta strand"
FT   Region          128..130
FT                   /note= "beta strand"
FT   Region          136..139
FT                   /note= "beta strand"
FT   Region          143..146
FT                   /note= "beta strand"
FT   Region          164..171
FT                   /note= "beta strand"
FT   Region          176..182
FT                   /note= "beta strand"
FT   Region          187..200
FT                   /note= "beta strand"
FT   Region          204..210
FT                   /note= "beta strand"
FT   Region          226..231
FT                   /note= "beta strand"
FT   Domain          111..135
FT                   /note= "receptor binding domain"
FT   Domain          170..174
FT                   /note= "receptor binding domain"
XX
PN   WO9910492-A1.
XX
PD   04-MAR-1999.
XX
PF   26-AUG-1998;   98WO-US17724.
XX
PR   26-AUG-1997;   97US-0056983.
XX
PA   (ZYMO ) ZYMOGENETICS INC.
XX
PI   Humes JM,  Sheppard PO;
XX
DR   WPI; 1999-204665/17.
DR   N-PSDB; AAX24684.
XX
PT   Zsig39 protein - used to modulate fatty acid metabolism
XX
PS   Claim 1; 111-112; 132pp; English.
XX
CC   This polypeptide comprises human adipocyte-specific protein zsig39,
CC   a protein that modulates free fatty acid metabolism.  zsig39 is a
CC   member of a family of proteins having a globular domain and a
CC   collagen-like domain capable of dimerisation or oligomerisation.
CC   zsig39 polypeptides were initially identified by querying an EST
CC   database for secretory signal sequences characterised by an upstream
CC   methionine start site, a hydrophobic region of approximately 13
CC   amino acids and a cleavage site.  A single EST sequence was
CC   discovered, and the novel polypeptide encoded by the full-length
CC   cDNA allowed the identification of a homologue relationship with
CC   adipocyte complement related protein Acrp30 and adipocyte secreted
CC   protein apM1.  A full-length clone (see AAX24684) was obtained from a
CC   lung tissue library.  Expression vectors, cultured cells and a
CC   method of producing zsig39 polypeptide are claimed, as well as
CC   zsig39 polypeptides having N- or C-terminal affinity tags, toxins,
CC   radionucleotides, enzymes or fluorophores, fusion proteins
CC   including zsig39 polypeptides, an antibody that specifically
CC   binds to an epitope of zsig39, and a method for modulating free
CC   fatty acid metabolism by administering a zsig39 polypeptide.  The
CC   zsig39 polypeptide may also be used in organ preservation, for
CC   cryopreservation, for surgical pretreatment to prevent injury due
CC   to ischaemia and/or inflammation, and as an antimicrobial agent,
CC   promoting lysis or phagocytosis of infectious agents.
XX
SQ   Sequence   243 AA;
```

```
  Query Match             100.0%;  Score 1325;  DB 20;  Length 243;
  Best Local Similarity   100.0%;  Pred. No. 5.3e-103;
  Matches  243;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MRPLLVLLLLGLAAGSPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MRPLLVLLLLGLAAGSPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPG 60

Qy         61 APGEKGEGGRPGLPGPRGDPGPRGEAGPAGPTGPAGECSVPPRSAFSAKRSESRVPPPSD 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 APGEKGEGGRPGLPGPRGDPGPRGEAGPAGPTGPAGECSVPPRSAFSAKRSESRVPPPSD 120

Qy        121 APLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGESIASFFQ 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 APLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGESIASFFQ 180

Qy        181 FFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWHSSP 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 FFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWHSSP 240

Qy        241 VFA 243
              |||
Db        241 VFA 243
```

```
RESULT 4
AAB33461
ID   AAB33461 standard; Protein; 243 AA.
XX
AC   AAB33461;
XX
DT   29-JAN-2001  (first entry)
XX
DE   Human PRO344 protein UNQ303 SEQ ID NO:241.
XX
KW   Human; immune related disease; diagnosis; antiinflammatory; cardiant;
KW   dermatological; antiarthritic; antirheumatic; immunosuppressive;
KW   haemostatic; antithyroid; antidiabetic; nootropic; neuroprotective;
KW   antianaemic; hepatotropic; virucide; antipsoriatic; antiallergic;
KW   antiasthmatic; systemic lupus erythematosus; rheumatoid arthritis;
KW   osteoarthritis; spondyloarthropathy; systemic sclerosis; sarcoidosis;
KW   idiopathic inflammatory myopathy; Sjogren's syndrome; thyroiditis;
KW   systemic vasculitis; autoimmune haemolytic anaemia; diabetes mellitus;
KW   autoimmune thrombocytopaenia; immune-mediated renal disease;
KW   demyelinating disease; hepatobiliary disease; Whipple's disease;
KW   inflammatory bowel disease; gluten-sensitive enteropathy;
KW   autoimmune disease; immune-mediated skin disease; allergic disease;
KW   immunological disease; transplantation associated disease;
KW   graft rejection; graft-versus-host-disease.
XX
OS   Homo sapiens.
XX
PN   WO200053758-A2.
```

XX
PD 14-SEP-2000.
XX
PF 02-MAR-2000; 2000WO-US05841.
XX
PR 08-MAR-1999; 99WO-US05028.
PR 10-MAR-1999; 99US-0123618.
PR 12-MAR-1999; 99US-0123957.
PR 23-MAR-1999; 99US-0125775.
PR 12-APR-1999; 99US-0128849.
PR 20-APR-1999; 99WO-US08615.
PR 28-APR-1999; 99US-0131445.
PR 04-MAY-1999; 99US-0132371.
PR 14-MAY-1999; 99US-0134287.
PR 02-JUN-1999; 99WO-US12252.
PR 23-JUN-1999; 99US-0141037.
PR 20-JUL-1999; 99US-0144758.
PR 26-JUL-1999; 99US-0145698.
PR 28-JUL-1999; 99US-0146222.
PR 01-SEP-1999; 99WO-US20111.
PR 08-SEP-1999; 99WO-US20594.
PR 13-SEP-1999; 99WO-US20944.
PR 15-SEP-1999; 99WO-US21090.
PR 15-SEP-1999; 99WO-US21547.
PR 05-OCT-1999; 99WO-US23089.
PR 29-OCT-1999; 99US-0162506.
PR 29-NOV-1999; 99WO-US28214.
PR 30-NOV-1999; 99WO-US28313.
PR 30-NOV-1999; 99WO-US28409.
PR 01-DEC-1999; 99WO-US28301.
PR 01-DEC-1999; 99WO-US28634.
PR 02-DEC-1999; 99WO-US28551.
PR 02-DEC-1999; 99WO-US28564.
PR 02-DEC-1999; 99WO-US28565.
PR 16-DEC-1999; 99WO-US30095.
PR 20-DEC-1999; 99WO-US30999.
PR 30-DEC-1999; 99WO-US31274.
PR 05-JAN-2000; 2000WO-US00219.
PR 06-JAN-2000; 2000WO-US00277.
PR 06-JAN-2000; 2000WO-US00376.
PR 11-FEB-2000; 2000WO-US03565.
PR 18-FEB-2000; 2000WO-US04341.
PR 18-FEB-2000; 2000WO-US04342.
PR 22-FEB-2000; 2000WO-US04414.
XX
PA (GETH ) GENENTECH INC.
XX
PI Ashkenazi AJ, Baker KP, Goddard A, Gurney AL, Hebert C, Henzel W;
PI Kabakoff RC, Lu Y, Pan J, Pennica D, Shelton DL, Smith V;
PI Stewart TA, Tumas D, Watanabe CK, Wood WI, Yan M;
XX
DR WPI; 2000-572271/53.
DR N-PSDB; AAC58626.
XX
PT Sixty four PRO polypeptides, useful in the diagnosis and treatment of
PT immune related disorders, e.g. systemic lupus erythematosis, rheumatoid
PT arthritis, osteoarthritis, thyroiditis and diabetes mellitus -
XX
PS Claim 33; Fig 96; 309pp; English.
XX
CC The present invention describes sixty four human PRO proteins which can
CC be used in the treatment of immune related diseases. The human PRO
CC proteins, anti-PRO antibodies, agonists and antagonists are useful for
CC treating and diagnosing immune related disorders. The disorders are
CC selected from systemic lupus erythematosus, rheumatoid arthritis,
CC osteoarthritis, juvenile chronic arthritis, spondyloarthropathies,
CC systemic sclerosis, idiopathic inflammatory myopathies, Sjogren's
CC syndrome, systemic vasculitis, sarcoidosis, autoimmune haemolytic
CC anaemia, autoimmune thrombocytopaenia, thyroiditis, diabetes mellitus,
CC immune-mediated renal disease, demyelinating diseases of the central
CC and peripheral nervous systems, hepatobiliary diseases, inflammatory
CC bowel disease, gluten-sensitive enteropathy and Whipple's disease,
CC autoimmune or immune-mediated skin diseases, allergic diseases,
CC immunological diseases of the lung, and transplantation associated
CC diseases including graft rejection and graft-versus-host-disease.
CC AAC58397 to AAC58578 represent PCR primers and hybridisation probes used
CC in the isolation of human PRO sequences. AAC58579 to AAC58642 and
CC AAB33414 to AAB33477 represent human PRO polynucleotide and protein
CC sequences given in the exemplification of the present invention.
XX
SQ Sequence 243 AA;

Query Match 100.0%; Score 1325; DB 21; Length 243;
Best Local Similarity 100.0%; Pred. No. 5.3e-103;
Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MRPLLVLLLLGLAAGSPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MRPLLVLLLLGLAAGSPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPG 60

Qy         61 APGEKGEGGRPGLPGPRGDPGPRGEAGPAGPTGPAGECSVPPRSAFSAKRSESRVPPPSD 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 APGEKGEGGRPGLPGPRGDPGPRGEAGPAGPTGPAGECSVPPRSAFSAKRSESRVPPPSD 120

Qy        121 APLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGESIASFFQ 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 APLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGESIASFFQ 180

Qy        181 FFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWHSSP 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 FFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWHSSP 240

Qy        241 VFA 243
              |||
Db        241 VFA 243
```

RESULT 5
AAY71468
ID AAY71468 standard; Protein; 243 AA.
XX
AC AAY71468;
XX
DT 08-NOV-2000 (first entry)
XX
DE Human PRO344 protein.
XX
KW PRO344; DNA40592-1242; human; ATCC No: 209492; antiproliferative;
KW neoplastic cell growth inhibitor; cytostatic; treatment; cancer; tumour;
KW breast; prostate; colon; lung; renal; ovarian; central nervous system;
KW CNS; leukemia; melanoma; Expressed Sequence Tag; EST; secreted protein;
KW extracellular domain; ECD.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Peptide 1..15
FT /label= Signal_peptide
FT Modified-site 11..17
FT /note= "N-myristoylation site"
FT Protein 16..243
FT /label= Mature_PRO344_protein
FT Modified-site 68..74
FT /note= "N-myristoylation site"
FT Binding-site 77..80
FT /note= "Cell attachment sequence"
FT Modified-site 216..222
FT /note= "N-myristoylation site"
XX
PN WO200032778-A2.
XX
PD 08-JUN-2000.
XX
PF 30-NOV-1999; 99WO-US28409.
XX
PR 01-DEC-1998; 98WO-US25108.

```
PR   16-DEC-1998;   98US-0112850.
PR   22-DEC-1998;   98US-0113296.
PR   20-JUL-1999;   99US-0144758.
PR   26-JUL-1999;   99US-0145698.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Chen J,  Goddard A,  Gurney AL,  Hillan K,  Napier M,  Wood WI;
XX
DR   WPI; 2000-412325/35.
DR   N-PSDB; AAD01241.
XX
PT   New composition useful for inhibiting neoplastic cell growth and for
PT   treating cancers, comprises PRO655, PRO344 or PRO364 polypeptide or
PT   their antagonists -
XX
PS   Claim 31; Fig 6; 108pp; English.
XX
CC   The present sequence is the human PRO344 protein, encoded by the cDNA
CC   clone, designated as DNA40592-1242. It is isolated from human foetal
CC   lung tissue, cDNA library, identified using probes based on a consensus
CC   sequence DNA34398, derived from secreted protein extracellular domain
CC   (ECD) expressed sequence tag (EST). This clone is assigned ATCC deposit
CC   No: 209492. PRO344 functions as a neoplastic cell growth inhibitor and
CC   is used for treating tumours, using an effective amount of PRO655, PRO364
CC   and PRO344. This composition is especially useful for treatment of human
CC   cancers such as breast, prostate, colon, lung, renal, ovarian and CNS,
CC   leukemia and melanoma.
XX
SQ   Sequence   243 AA;

  Query Match             100.0%;  Score 1325;  DB 21;  Length 243;
  Best Local Similarity   100.0%;  Pred. No. 5.3e-103;
  Matches  243;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MRPLLVLLLLGLAAGSPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MRPLLVLLLLGLAAGSPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPG 60

Qy         61 APGEKGEGGRPGLPGPRGDPGPRGEAGPAGPTGPAGECSVPPRSAFSAKRSESRVPPPSD 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 APGEKGEGGRPGLPGPRGDPGPRGEAGPAGPTGPAGECSVPPRSAFSAKRSESRVPPPSD 120

Qy        121 APLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGESIASFFQ 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 APLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGESIASFFQ 180

Qy        181 FFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWHSSP 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 FFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWHSSP 240

Qy        241 VFA 243
              |||
Db        241 VFA 243


RESULT 6
AAY93688
ID   AAY93688 standard; Protein; 243 AA.
XX
AC   AAY93688;
XX
DT   03-OCT-2000  (first entry)
XX
DE   Amino acid sequence of novel polypeptide PRO344.
XX
KW   PRO201; PRO292; PRO327; PRO1265;  PRO344; PRO343; PRO347; PRO357;
KW   PRO715; PRO1017; PRO1112; PRO509; PRO853; PRO882; tumour cell;
KW   tumourigenesis; cancer; neoplastic cell growth; cell proliferation.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1..15
FT                   /note= "signal peptide"
FT   Modified-site   11..17
FT                   /note= "N-myristoylation site"
FT   Modified-site   68..74
FT                   /note= "N-myristoylation site"
FT   Region          77..80
FT                   /note= "cell attachment sequence"
FT   Modified-site   216..222
FT                   /note= "N-myristoylation site"
XX
PN   WO200037640-A2.
XX
PD   29-JUN-2000.
XX
PF   16-DEC-1999;   99WO-US30095.
XX
PR   22-DEC-1998;   98US-0113296.
PR   08-MAR-1999;   99WO-US05028.
PR   02-JUN-1999;   99WO-US12252.
PR   01-SEP-1999;   99WO-US20111.
PR   15-SEP-1999;   99WO-US21090.
PR   30-NOV-1999;   99WO-US28313.
PR   30-NOV-1999;   99WO-US28409.
PR   01-DEC-1999;   99WO-US28301.
PR   02-DEC-1999;   99WO-US28565.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Botstein D,  Goddard A,  Gurney AL,  Hillan K,  Lawrence DA,  Roy MA;
PI   Wood WI;
XX
DR   WPI; 2000-452188/39.
DR   N-PSDB; AAA46907.
XX
PT   New anti-polypeptide antibody useful in the treatment and diagnosis of
PT   neoplastic cell growth and proliferation -
XX
PS   Claim 61; Fig 10; 220pp; English.
XX
CC   The present sequence represents a novel human polypeptide. The
CC   specification describes novel polypeptides designated PRO201, PRO292,
CC   PRO327, PRO1265, PRO344, PRO343, PRO347, PRO357, PRO715, PRO1017,
CC   PRO1112, PRO509, PRO853 and PRO882. These genes are amplified in
CC   the genome of tumour cells. The polypeptides are believed to contribute
CC   to tumourigenesis. The polypeptides are useful target for the
CC   identification of certain cancers, and may act as predictors of the
CC   prognosis of tumour treatment. Antibodies against these polypeptides
CC   are useful in the treatment and diagnosis of neoplastic cell growth
CC   and proliferation in mammals.
XX
SQ   Sequence   243 AA;

  Query Match             100.0%;  Score 1325;  DB 21;  Length 243;
  Best Local Similarity   100.0%;  Pred. No. 5.3e-103;
  Matches  243;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MRPLLVLLLLGLAAGSPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MRPLLVLLLLGLAAGSPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPG 60

Qy         61 APGEKGEGGRPGLPGPRGDPGPRGEAGPAGPTGPAGECSVPPRSAFSAKRSESRVPPPSD 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 APGEKGEGGRPGLPGPRGDPGPRGEAGPAGPTGPAGECSVPPRSAFSAKRSESRVPPPSD 120

Qy        121 APLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGESIASFFQ 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 APLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGESIASFFQ 180

Qy        181 FFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWHSSP 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 FFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWHSSP 240
```

```
Qy        241 VFA 243
              |||
Db        241 VFA 243


RESULT 7
AAB01318
ID   AAB01318 standard; Protein; 243 AA.
XX
AC   AAB01318;
XX
DT   25-SEP-2000  (first entry)
XX
DE   Human PRO344 polypeptide.
XX
KW   PRO; membrane bound protein; secreted protein; PRO357; PRO327;
KW   PRO243; PRO715; PRO241; PRO323; PRO299; PRO233; PRO344; PRO347;
KW   PRO355; PRO353; PRO361; PRO365; transmembrane polypeptide;
KW   antibody; screening; detection; inhibition; probe; primer; human.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1..15
FT                   /label= Signal peptide
FT   Modified-site   11..17
FT                   /note= "N-myristoylation site"
FT   Modified-site   68..74
FT                   /note= "N-myristoylation site"
FT   Region          77..80
FT                   /label= Cell attachment sequence
FT   Modified-site   216..222
FT                   /note= "N-myristoylation site"
XX
PN   WO200032776-A2.
XX
PD   08-JUN-2000.
XX
PF   01-DEC-1999;   99WO-US28301.
XX
PR   01-DEC-1998;   98WO-US25108.
PR   16-DEC-1998;   98US-0112850.
PR   22-DEC-1998;   98US-0113296.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Baker KP,  Botstein D,  Eaton DL,  Ferrara N,  Filvaroff E;
PI   Gerritsen ME,  Goddard A,  Godowski PJ,  Grimaldi CJ,  Gurney AL;
PI   Hillan KJ,  Kljavin IJ,  Napier MA,  Roy MA,  Tumas D,  Wood WI;
XX
DR   WPI; 2000-412324/35.
DR   N-PSDB; AAA49560.
XX
PT   New human nucleic acids encoding secreted and transmembrane
PT   polypeptides, designated as PRO polypeptides, useful as pharmaceutical
PT   and diagnostic agents
XX
PS   Claim 12; Fig 18; 187pp; English.
XX
CC   New human nucleic acids encoding secreted and transmembrane
CC   polypeptides which are designated as PRO polypeptides are described
CC   The membrane-bound proteins have various industrial applications,
CC   including as pharmaceutical and diagnostic agents. The membrane-bound
CC   proteins can also be employed for screening of potential peptide or
CC   small molecule inhibitors of the relevant receptor/ligand interaction.
CC   Anti-PRO antibodies are useful for the affinity purification of PRO
CC   from recombinant cell culture or natural sources.
XX
SQ   Sequence   243 AA;

  Query Match           100.0%;  Score 1325;  DB 21;  Length 243;
  Best Local Similarity  100.0%;  Pred. No. 5.3e-103;
  Matches  243;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MRPLLVLLLLGLAAGSPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MRPLLVLLLLGLAAGSPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPG 60

Qy         61 APGEKGEGGRPGLPGPRGDPGPRGEAGPAGPTGPAGECSVPPRSAFSAKRSESRVPPPSD 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 APGEKGEGGRPGLPGPRGDPGPRGEAGPAGPTGPAGECSVPPRSAFSAKRSESRVPPPSD 120

Qy        121 APLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGESIASFFQ 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 APLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGESIASFFQ 180

Qy        181 FFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWHSSP 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 FFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWHSSP 240

Qy        241 VFA 243
              |||
Db        241 VFA 243


RESULT 8
AAU12352
ID   AAU12352 standard; Protein; 243 AA.
XX
AC   AAU12352;
XX
DT   24-OCT-2001  (first entry)
XX
DE   Human PRO344 polypeptide sequence.
XX
KW   Human secretory and transmembrane; PRO; mammalian; cancer; lung;
KW   breast; prostate; cervical; tumour necrosis factor-alpha; TNF-alpha;
KW   cartilage; ear; proliferation; glucose; free fatty acid; skeletal muscle;
KW   adipocyte; A-peptide; factor VIIA; gene therapy.
XX
OS   Homo sapiens.
XX
PN   WO200140466-A2.
XX
PD   07-JUN-2001.
XX
PF   01-DEC-2000; 2000WO-US32678.
XX
PR   01-DEC-1999;   99WO-US28301.
PR   01-DEC-1999;   99WO-US28634.
PR   02-DEC-1999;   99WO-US28551.
PR   02-DEC-1999;   99WO-US28564.
PR   02-DEC-1999;   99WO-US28565.
PR   09-DEC-1999;   99US-0170262.
PR   16-DEC-1999;   99WO-US30095.
PR   20-DEC-1999;   99WO-US30911.
PR   20-DEC-1999;   99WO-US30999.
PR   30-DEC-1999;   99WO-US31243.
PR   06-JAN-2000; 2000WO-US00277.
PR   06-JAN-2000; 2000WO-US00376.
PR   11-FEB-2000; 2000WO-US03565.
PR   18-FEB-2000; 2000WO-US04341.
PR   18-FEB-2000; 2000WO-US04342.
PR   22-FEB-2000; 2000WO-US04414.
PR   24-FEB-2000; 2000WO-US04914.
PR   24-FEB-2000; 2000WO-US05004.
PR   01-MAR-2000; 2000WO-US05601.
PR   20-MAR-2000; 2000WO-US07377.
PR   21-MAR-2000; 2000WO-US07532.
PR   30-MAR-2000; 2000WO-US08439.
PR   17-MAY-2000; 2000WO-US13705.
PR   22-MAY-2000; 2000WO-US14042.
PR   30-MAY-2000; 2000WO-US14941.
PR   02-JUN-2000; 2000WO-US15264.
PR   10-NOV-2000; 2000WO-US30873.
XX
```

```
PA   (GETH ) GENENTECH INC.
XX
PI   Baker KP,  Beresini M,  Deforge L,  Desnoyers L,  Filvaroff E,  Gao W;
PI   Gerritsen ME,  Goddard A,  Godowski PJ,  Gurney AL,  Sherwood S;
PI   Smith V,  Stewart TA,  Tumas D,  Watanabe CK,  Wood WI,  Zhang Z;
XX
DR   WPI; 2001-408281/43.
DR   N-PSDB; AAS21424.
XX
PT   Isolated, secretory and transmembrane PRO polypeptide used to detect
PT   other PRO polypeptides, link bioactive molecules to cells expressing
PT   PRO polypeptides, and detect the presence of mammalian tumours e.g.
PT   lung, breast, prostate, cervical -
XX
PS   Claim 12; Fig 362; 813pp; English.
XX
CC   AAU12172-AAU12446 represent novel human secretory and transmembrane
CC   PRO polypeptides. The PRO polypeptides are useful to detect other
CC   PRO polypeptides, to link bioactive molecules to cells expressing
CC   PRO polypeptides, to modulate biological activities of cells expressing
CC   PRO polypeptides, and to detect the presence of mammalian lung, colon,
CC   breast, prostate, rectal, cervical or liver tumours by comparing PRO
CC   polypeptide expression in a cell sample to that in a control sample.
CC   Some of the 275 sequences are also useful to stimulate the release of
CC   tumour necrosis factor-alpha (TNF-alpha) from human blood, the
CC   proliferation or differentiation of chondrocytes, the proliferation or
CC   gene expression in pericyte cells, the release of proteoglycans from
CC   cartilage, the proliferation of inner ear utricular supporting cells or
CC   of T-lymphocytes, the release of a cytokine from peripheral blood
CC   monocytes (PBMCs), or the proliferation of endothelial cells. Some of
CC   the PRO polypeptides may modulate glucose or free fatty acid uptake by
CC   skeletal muscle cells or by adipocytes; or inhibit binding of A-peptide
CC   to factor VIIA. The PRO polypeptides can be used in assays to identify
CC   molecules involved in binding interactions. The polynucleotides encoding
CC   PRO polypeptides can be used to generate probes, antisense RNA/DNA,
CC   transgenic or knock out animals and can be used in gene therapy.
XX
SQ   Sequence   243 AA;

  Query Match             100.0%;  Score 1325;  DB 22;  Length 243;
  Best Local Similarity   100.0%;  Pred. No. 5.3e-103;
  Matches  243;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MRPLLVLLLLGLAAGSPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MRPLLVLLLLGLAAGSPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPG 60

Qy         61 APGEKGEGGRPGLPGPRGDPGPRGEAGPAGPTGPAGECSVPPRSAFSAKRSESRVPPPSD 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 APGEKGEGGRPGLPGPRGDPGPRGEAGPAGPTGPAGECSVPPRSAFSAKRSESRVPPPSD 120

Qy        121 APLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGESIASFFQ 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 APLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGESIASFFQ 180

Qy        181 FFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWHSSP 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 FFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWHSSP 240

Qy        241 VFA 243
              |||
Db        241 VFA 243


RESULT 9
AAB65815
ID   AAB65815 standard; Protein; 243 AA.
XX
AC   AAB65815;
XX
DT   28-MAR-2001  (first entry)
XX
DE   Human TANGO 253 SEQ ID NO: 3.
XX
KW   Human; mouse; secreted protein; TANGO253; TANGO 257; TANGO 281;
KW   INTERCEPT 258; coronary disorder; olfactory disorder;
KW   neurological disorder; pulmonary disorder; immunological disorder;
KW   developmental disorder; kidney disorder.
XX
OS   Homo sapiens.
XX
PN   WO200078808-A1.
XX
PD   28-DEC-2000.
XX
PF   19-JUN-2000; 2000WO-US16883.
XX
PR   18-JUN-1999;   99US-0336536.
XX
PA   (MILL-) MILLENNIUM PHARM INC.
XX
PI   Leiby KR,  McKay C,  Bossone S;
XX
DR   WPI; 2001-050109/06.
XX
PT   New nucleic acids for treating diseases and disorders, e.g.
PT   atherosclerosis, infection, autoimmune diseases, obesity, ear
PT   disorders, brain disorders, tumors, diabetes, arthritis, multiple
PT   sclerosis and asthma -
XX
PS   Claim 9; Page 211-212; 332pp; English.
XX
CC   The present invention provides the protein and coding sequences of the
CC   human and murine secreted or transmembrane proteins TANGO 253, TANGO 257,
CC   TANGO 281 and INTERCEPT 258. These are useful in the treatment of
CC   coronary, pulmonary, olfactory, immunological, neurological,
CC   developmental and kidney disorders.
XX
SQ   Sequence   243 AA;

  Query Match             100.0%;  Score 1325;  DB 22;  Length 243;
  Best Local Similarity   100.0%;  Pred. No. 5.3e-103;
  Matches  243;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MRPLLVLLLLGLAAGSPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MRPLLVLLLLGLAAGSPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPG 60

Qy         61 APGEKGEGGRPGLPGPRGDPGPRGEAGPAGPTGPAGECSVPPRSAFSAKRSESRVPPPSD 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 APGEKGEGGRPGLPGPRGDPGPRGEAGPAGPTGPAGECSVPPRSAFSAKRSESRVPPPSD 120

Qy        121 APLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGESIASFFQ 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 APLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGESIASFFQ 180

Qy        181 FFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWHSSP 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 FFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWHSSP 240

Qy        241 VFA 243
              |||
Db        241 VFA 243


RESULT 10
AAB49593
ID   AAB49593 standard; Protein; 243 AA.
XX
AC   AAB49593;
XX
DT   13-MAR-2001  (first entry)
XX
DE   Human adipocyte complement related protein homolog zsig39.
XX
KW   Human; zacrp5; gene therapy; complement inhibition; C1q domain;
```

KW adipocyte complement related protein homolog;
KW inflammation; hormone secretion; inositol phosphate; arachidonate;
KW phospholipase C activation; gastric emptying; neutrophil activation;
KW superoxide anion production; antimicrobial; acute vascular injury;
KW wound healing; zsig39.
XX
OS Homo sapiens.
XX
PN WO200073444-A1.
XX
PD 07-DEC-2000.
XX
PF 18-MAY-2000; 2000WO-US13608.
XX
PR 27-MAY-1999; 99US-0321372.
XX
PA (ZYMO ) ZYMOGENETICS INC.
XX
PI Piddington CS, Sheppard PO;
XX
DR WPI; 2001-061531/07.
XX
PT Novel adipocyte complement related protein homolog, ZACRP5, useful for
PT diagnosing and treating inflammation, vascular injury microbial
PT infections, and in wound healing -
XX
PS Disclosure; Fig 1; 121pp; English.
XX
CC The present invention relates to human adipocyte complement related
CC protein homolog, zacrp5 protein and coding sequence (see AAB49590 and
CC AAC90045). ZACRP5 has a carboxyl-terminal C1q domain. The zacrp5 gene is
CC located on human chromosome 16. zacrp5 gene and protein are useful for
CC diagnosing and treating inflammations, for determining arterial
CC remodelling, for modulating calcium ion concentration, hormone
CC secretion, DNA synthesis or cell growth, inositol phosphate turnover,
CC arachidonate release, phospholipase C activation, gastric emptying, human
CC neutrophil activation or ADCC capability and superoxide anion production.
CC zacrp5 gene and protein are also useful as antimicrobial applications,
CC preferably against bacteria and virus, for complement inhibition, for
CC treating acute vascular injury, and for wound healing. The present
CC sequence is human adipocyte complement related protein homolog, zsig39
CC protein. This protein was used in a sequence homology alignment with
CC zacrp5 protein.
XX
SQ Sequence 243 AA;

Query Match 100.0%; Score 1325; DB 22; Length 243;
Best Local Similarity 100.0%; Pred. No. 5.3e-103;
Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MRPLLVLLLLGLAAGSPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MRPLLVLLLLGLAAGSPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPG 60

Qy         61 APGEKGEGGRPGLPGPRGDPGPRGEAGPAGPTGPAGECSVPPRSAFSAKRSESRVPPPSD 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 APGEKGEGGRPGLPGPRGDPGPRGEAGPAGPTGPAGECSVPPRSAFSAKRSESRVPPPSD 120

Qy        121 APLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGESIASFFQ 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 APLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGESIASFFQ 180

Qy        181 FFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWHSSP 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 FFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWHSSP 240

Qy        241 VFA 243
              |||
Db        241 VFA 243
```

RESULT 11
AAB49599

ID AAB49599 standard; Protein; 243 AA.
XX
AC AAB49599;
XX
DT 13-MAR-2001 (first entry)
XX
DE Human adipocyte complement related protein homolog zsig39.
XX
KW Human; zacrp6; gene therapy; complement inhibition; C1q domain;
KW adipocyte complement related protein homolog;
KW inflammation; hormone secretion; inositol phosphate; arachidonate;
KW phospholipase C activation; gastric emptying; neutrophil activation;
KW superoxide anion production; antimicrobial; acute vascular injury;
KW wound healing; zsig39.
XX
OS Homo sapiens.
XX
PN WO200073446-A2.
XX
PD 07-DEC-2000.
XX
PF 22-MAY-2000; 2000WO-US14024.
XX
PR 27-MAY-1999; 99US-0321262.
XX
PA (ZYMO ) ZYMOGENETICS INC.
XX
PI Piddington CS, Sheppard PO;
XX
DR WPI; 2001-061532/07.
XX
PT Novel adipocyte complement related protein homolog, ZACRP6, useful as
PT modulators of neurotransmission and for treating disseminated
PT intravascular coagulation, arteriosclerosis and acute vascular injury
PT -
XX
PS Disclosure; Fig 1; 119pp; English.
XX
CC The present invention relates to human adipocyte complement related
CC protein homolog, ZACRP6 protein and coding sequence (see AAB49596 and
CC AAC90051). ZACRP6 has a carboxyl-terminal C1q domain. The zacrp6 gene is
CC located on human chromosome 21q. zacrp6 gene and protein are useful for
CC diagnosing and treating inflammations, for determining arterial
CC remodelling, for modulating calcium ion concentration, hormone
CC secretion, DNA synthesis or cell growth, inositol phosphate turnover,
CC arachidonate release, phospholipase C activation, gastric emptying, human
CC neutrophil activation or ADCC capability and superoxide anion production.
CC zacrp6 gene and protein are also useful as antimicrobial applications,
CC preferably against bacteria and virus, for complement inhibition, for
CC treating acute vascular injury, disseminated intravascular coagulation,
CC arteriosclerosis and for wound healing. The present sequence is human
CC adipocyte complement related protein homolog zsig39. This protein was
CC used in a sequence homology comparison with ZACRP6 protein.
XX
SQ Sequence 243 AA;

Query Match 100.0%; Score 1325; DB 22; Length 243;
Best Local Similarity 100.0%; Pred. No. 5.3e-103;
Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MRPLLVLLLLGLAAGSPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MRPLLVLLLLGLAAGSPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPG 60

Qy         61 APGEKGEGGRPGLPGPRGDPGPRGEAGPAGPTGPAGECSVPPRSAFSAKRSESRVPPPSD 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 APGEKGEGGRPGLPGPRGDPGPRGEAGPAGPTGPAGECSVPPRSAFSAKRSESRVPPPSD 120

Qy        121 APLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGESIASFFQ 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 APLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGESIASFFQ 180

Qy        181 FFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWHSSP 240
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 FFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWHSSP 240

Qy        241 VFA 243
                |||
Db        241 VFA 243


RESULT 12
AAB65891
ID   AAB65891 standard; Protein; 243 AA.
XX
AC   AAB65891;
XX
DT   28-MAR-2001  (first entry)
XX
DE   Human secreted protein related protein SEQ ID NO: 108.
XX
KW   Human; mouse; secreted protein; TANGO253; TANGO 257; TANGO 281;
KW   INTERCEPT 258; coronary disorder; olfactory disorder;
KW   neurological disorder; pulmonary disorder; immunological disorder;
KW   developmental disorder; kidney disorder.
XX
OS   Homo sapiens.
XX
PN   WO200078808-A1.
XX
PD   28-DEC-2000.
XX
PF   19-JUN-2000; 2000WO-US16883.
XX
PR   18-JUN-1999;   99US-0336536.
XX
PA   (MILL-) MILLENNIUM PHARM INC.
XX
PI   Leiby KR,  McKay C,  Bossone S;
XX
DR   WPI; 2001-050109/06.
XX
PT   New nucleic acids for treating diseases and disorders, e.g.
PT   atherosclerosis, infection, autoimmune diseases, obesity, ear
PT   disorders, brain disorders, tumors, diabetes, arthritis, multiple
PT   sclerosis and asthma -
XX
PS   Disclosure; Page 274; 332pp; English.
XX
CC   The present invention provides the protein and coding sequences of the
CC   human and murine secreted or transmembrane proteins TANGO 253, TANGO 257,
CC   TANGO 281 and INTERCEPT 258. These are useful in the treatment of
CC   coronary, pulmonary, olfactory, immunological, neurological,
CC   developmental and kidney disorders.
XX
SQ   Sequence   243 AA;

  Query Match             99.8%;  Score 1323;  DB 22;  Length 243;
  Best Local Similarity   99.6%;  Pred. No. 7.8e-103;
  Matches  242;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MRPLLVLLLLGLAAGSPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MRPLLVLLLLGLAAGSPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPG 60

Qy         61 APGEKGEGGRPGLPGPRGDPGPRGEAGPAGPTGPAGECSVPPRSAFSAKRSESRVPPPSD 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 APGEKGEGGRPGLPGPRGDPGPRGEAGPAGPTGPAGECSVPPRSAFSAKRSESRVPPPSD 120

Qy        121 APLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGESIASFFQ 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||:||||||||
Db        121 APLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGESLASFFQ 180

Qy        181 FFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWHSSP 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 FFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWHSSP 240

Qy        241 VFA 243
                |||
Db        241 VFA 243


RESULT 13
AAB65888
ID   AAB65888 standard; Protein; 243 AA.
XX
AC   AAB65888;
XX
DT   28-MAR-2001  (first entry)
XX
DE   Human secreted protein related protein SEQ ID NO: 102.
XX
KW   Human; mouse; secreted protein; TANGO253; TANGO 257; TANGO 281;
KW   INTERCEPT 258; coronary disorder; olfactory disorder;
KW   neurological disorder; pulmonary disorder; immunological disorder;
KW   developmental disorder; kidney disorder.
XX
OS   Homo sapiens.
XX
PN   WO200078808-A1.
XX
PD   28-DEC-2000.
XX
PF   19-JUN-2000; 2000WO-US16883.
XX
PR   18-JUN-1999;   99US-0336536.
XX
PA   (MILL-) MILLENNIUM PHARM INC.
XX
PI   Leiby KR,  McKay C,  Bossone S;
XX
DR   WPI; 2001-050109/06.
XX
PT   New nucleic acids for treating diseases and disorders, e.g.
PT   atherosclerosis, infection, autoimmune diseases, obesity, ear
PT   disorders, brain disorders, tumors, diabetes, arthritis, multiple
PT   sclerosis and asthma -
XX
PS   Disclosure; Page 270-271; 332pp; English.
XX
CC   The present invention provides the protein and coding sequences of the
CC   human and murine secreted or transmembrane proteins TANGO 253, TANGO 257,
CC   TANGO 281 and INTERCEPT 258. These are useful in the treatment of
CC   coronary, pulmonary, olfactory, immunological, neurological,
CC   developmental and kidney disorders.
XX
SQ   Sequence   243 AA;

  Query Match             99.7%;  Score 1321;  DB 22;  Length 243;
  Best Local Similarity   99.6%;  Pred. No. 1.1e-102;
  Matches  242;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 MRPLLVLLLLGLAAGSPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPG 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db          1 MRPLLVLLLLGLAAGSPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGVPG 60

Qy         61 APGEKGEGGRPGLPGPRGDPGPRGEAGPAGPTGPAGECSVPPRSAFSAKRSESRVPPPSD 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 APGEKGEGGRPGLPGPRGDPGPRGEAGPAGPTGPAGECSVPPRSAFSAKRSESRVPPPSD 120

Qy        121 APLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGESIASFFQ 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 APLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGESIASFFQ 180

Qy        181 FFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWHSSP 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 FFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWHSSP 240

Qy        241 VFA 243
```

```
                    |||
Db        241 VFA 243


RESULT 14
AAB65889
ID   AAB65889 standard; Protein; 243 AA.
XX
AC   AAB65889;
XX
DT   28-MAR-2001  (first entry)
XX
DE   Human secreted protein related protein SEQ ID NO: 104.
XX
KW   Human; mouse; secreted protein; TANGO253; TANGO 257; TANGO 281;
KW   INTERCEPT 258; coronary disorder; olfactory disorder;
KW   neurological disorder; pulmonary disorder; immunological disorder;
KW   developmental disorder; kidney disorder.
XX
OS   Homo sapiens.
XX
PN   WO200078808-A1.
XX
PD   28-DEC-2000.
XX
PF   19-JUN-2000; 2000WO-US16883.
XX
PR   18-JUN-1999;   99US-0336536.
XX
PA   (MILL-) MILLENNIUM PHARM INC.
XX
PI   Leiby KR,  McKay C,  Bossone S;
XX
DR   WPI; 2001-050109/06.
XX
PT   New nucleic acids for treating diseases and disorders, e.g.
PT   atherosclerosis, infection, autoimmune diseases, obesity, ear
PT   disorders, brain disorders, tumors, diabetes, arthritis, multiple
PT   sclerosis and asthma -
XX
PS   Disclosure; Page 271-272; 332pp; English.
XX
CC   The present invention provides the protein and coding sequences of the
CC   human and murine secreted or transmembrane proteins TANGO 253, TANGO 257,
CC   TANGO 281 and INTERCEPT 258. These are useful in the treatment of
CC   coronary, pulmonary, olfactory, immunological, neurological,
CC   developmental and kidney disorders.
XX
SQ   Sequence   243 AA;

  Query Match           99.7%;  Score 1321;  DB 22;  Length 243;
  Best Local Similarity  99.6%;  Pred. No. 1.1e-102;
  Matches  242;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 MRPLLVLLLLGLAAGSPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MRPLLVLLLLGLAAGSPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPG 60

Qy         61 APGEKGEGGRPGLPGPRGDPGPRGEAGPAGPTGPAGECSVPPRSAFSAKRSESRVPPPSD 120
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db         61 APGEKGEGGRPGLPGPRGDPGPRGEAGPAGPTGPVGECSVPPRSAFSAKRSESRVPPPSD 120

Qy        121 APLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGESIASFFQ 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 APLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGESIASFFQ 180

Qy        181 FFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWHSSP 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 FFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWHSSP 240

Qy        241 VFA 243
              |||
Db        241 VFA 243


RESULT 15
AAB65890
ID   AAB65890 standard; Protein; 243 AA.
XX
AC   AAB65890;
XX
DT   28-MAR-2001  (first entry)
XX
DE   Human secreted protein related protein SEQ ID NO: 106.
XX
KW   Human; mouse; secreted protein; TANGO253; TANGO 257; TANGO 281;
KW   INTERCEPT 258; coronary disorder; olfactory disorder;
KW   neurological disorder; pulmonary disorder; immunological disorder;
KW   developmental disorder; kidney disorder.
XX
OS   Homo sapiens.
XX
PN   WO200078808-A1.
XX
PD   28-DEC-2000.
XX
PF   19-JUN-2000; 2000WO-US16883.
XX
PR   18-JUN-1999;   99US-0336536.
XX
PA   (MILL-) MILLENNIUM PHARM INC.
XX
PI   Leiby KR,  McKay C,  Bossone S;
XX
DR   WPI; 2001-050109/06.
XX
PT   New nucleic acids for treating diseases and disorders, e.g.
PT   atherosclerosis, infection, autoimmune diseases, obesity, ear
PT   disorders, brain disorders, tumors, diabetes, arthritis, multiple
PT   sclerosis and asthma -
XX
PS   Disclosure; Page 272-273; 332pp; English.
XX
CC   The present invention provides the protein and coding sequences of the
CC   human and murine secreted or transmembrane proteins TANGO 253, TANGO 257,
CC   TANGO 281 and INTERCEPT 258. These are useful in the treatment of
CC   coronary, pulmonary, olfactory, immunological, neurological,
CC   developmental and kidney disorders.
XX
SQ   Sequence   243 AA;

  Query Match           99.7%;  Score 1321;  DB 22;  Length 243;
  Best Local Similarity  99.6%;  Pred. No. 1.1e-102;
  Matches  242;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 MRPLLVLLLLGLAAGSPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MRPLLVLLLLGLAAGSPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPG 60

Qy         61 APGEKGEGGRPGLPGPRGDPGPRGEAGPAGPTGPAGECSVPPRSAFSAKRSESRVPPPSD 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 APGEKGEGGRPGLPGPRGDPGPRGEAGPAGPTGPAGECSVPPRSAFSAKRSESRVPPPSD 120

Qy        121 APLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGESIASFFQ 180
              |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 APLPFDRVLANEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGESIASFFQ 180

Qy        181 FFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWHSSP 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 FFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWHSSP 240

Qy        241 VFA 243
              |||
Db        241 VFA 243
```

Search completed: June 18, 2003, 14:51:05
Job time : 87 secs

GenCore version 5.1.6


OM protein - protein search, using sw model

Run on: June 18, 2003, 14:41:33 ; Search time 81 Seconds
(without alignments)
618.141 Million cell updates/sec

Title: US-09-943-851A-42
Perfect score: 1325
Sequence: 1 MRPLLVLLLLGLAAGSPPLD.....DSTFSGFLVYSDWHSSPVFA 243

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_21:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 1254 | 94.6 | 243 | 11 | Q8R002 | Q8r002 mus musculu |
| 2 | 425.5 | 32.1 | 240 | 6 | Q95MQ4 | Q95mq4 bos taurus |
| 3 | 422 | 31.8 | 243 | 6 | Q95JD7 | Q95jd7 macaca mula |
| 4 | 415.5 | 31.4 | 675 | 6 | Q9N178 | Q9n178 sus scrofa |
| 5 | 402 | 30.3 | 295 | 11 | Q9Z1K4 | Q9z1k4 rattus norv |
| 6 | 395 | 29.8 | 294 | 11 | Q9D8U4 | Q9d8u4 mus musculu |
| 7 | 373.5 | 28.2 | 194 | 6 | Q95J95 | Q95j95 canis famil |
| 8 | 370.5 | 28.0 | 744 | 11 | Q9D2V4 | Q9d2v4 mus musculu |
| 9 | 370.5 | 28.0 | 744 | 11 | Q921S8 | Q921s8 mus musculu |
| 10 | 368 | 27.8 | 744 | 4 | Q96D07 | Q96d07 homo sapien |
| 11 | 363.5 | 27.4 | 705 | 4 | Q8TEJ5 | Q8tej5 homo sapien |
| 12 | 314 | 23.7 | 245 | 11 | Q9DCM6 | Q9dcm6 mus musculu |
| 13 | 270.5 | 20.4 | 246 | 11 | Q9ES30 | Q9es30 mus musculu |
| 14 | 266.5 | 20.1 | 196 | 11 | Q920N0 | Q920n0 tamias sibi |
| 15 | 252.5 | 19.1 | 246 | 13 | Q9I9Q7 | Q9i9q7 carassius a |
| 16 | 247.5 | 18.7 | 256 | 13 | Q9I9Q9 | Q9i9q9 cyprinus ca |
| 17 | 241.5 | 18.2 | 347 | 4 | Q96IH6 | Q96ih6 homo sapien |
| 18 | 238.5 | 18.0 | 583 | 4 | Q96G58 | Q96g58 homo sapien |
| 19 | 238.5 | 18.0 | 992 | 4 | Q9UG76 | Q9ug76 homo sapien |
| 20 | 238.5 | 18.0 | 1016 | 4 | Q9Y6C2 | Q9y6c2 homo sapien |
| 21 | 237 | 17.9 | 890 | 5 | O77087 | O77087 alvinella p |
| 22 | 235 | 17.7 | 1017 | 11 | Q99K41 | Q99k41 mus musculu |
| 23 | 234.5 | 17.7 | 173 | 6 | O62789 | O62789 sus scrofa |
| 24 | 225 | 17.0 | 325 | 5 | O17036 | O17036 caenorhabdi |
| 25 | 223 | 16.8 | 281 | 11 | Q9QXP7 | Q9qxp7 mus musculu |
| 26 | 222.5 | 16.8 | 322 | 5 | O01945 | O01945 meloidogyne |
| 27 | 222 | 16.8 | 251 | 13 | Q9I9Q8 | Q9i9q8 brachydanio |
| 28 | 222 | 16.8 | 1378 | 5 | O97405 | O97405 haliotis di |
| 29 | 221.5 | 16.7 | 120 | 6 | O77782 | O77782 oryctolagus |
| 30 | 221.5 | 16.7 | 248 | 6 | Q9TT06 | Q9tt06 ovis aries |
| 31 | 220 | 16.6 | 341 | 5 | Q20142 | Q20142 caenorhabdi |
| 32 | 219.5 | 16.6 | 205 | 11 | Q9D0W2 | Q9d0w2 mus musculu |
| 33 | 219.5 | 16.6 | 326 | 11 | Q8R066 | Q8r066 mus musculu |
| 34 | 219.5 | 16.6 | 381 | 5 | Q94399 | Q94399 caenorhabdi |
| 35 | 219.5 | 16.6 | 589 | 11 | Q99LL6 | Q99ll6 mus musculu |
| 36 | 219.5 | 16.6 | 1453 | 11 | Q63079 | Q63079 rattus norv |
| 37 | 218 | 16.5 | 182 | 11 | Q8R1P2 | Q8r1p2 mus musculu |
| 38 | 217 | 16.4 | 319 | 5 | O17038 | O17038 caenorhabdi |
| 39 | 217 | 16.4 | 1160 | 4 | Q14046 | Q14046 homo sapien |
| 40 | 217 | 16.4 | 1344 | 13 | O93419 | O93419 gallus gall |
| 41 | 217 | 16.4 | 1418 | 6 | Q28396 | Q28396 equus cabal |
| 42 | 217 | 16.4 | 1442 | 11 | Q62031 | Q62031 mus musculu |
| 43 | 217 | 16.4 | 1442 | 11 | Q62033 | Q62033 mus musculu |
| 44 | 217 | 16.4 | 1450 | 13 | Q9YIB4 | Q9yib4 cynops pyrr |
| 45 | 217 | 16.4 | 1459 | 11 | Q62032 | Q62032 mus musculu |

ALIGNMENTS

```
RESULT 1
Q8R002
ID   Q8R002      PRELIMINARY;      PRT;   243 AA.
AC   Q8R002;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Similar to DKFZP586B0621 protein (Hypothetical 25.4 kDa protein).
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Strausberg R.;
RL   Submitted (FEB-2002) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Strausberg R.;
RL   Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC023068; AAH23068.1; -.
DR   EMBL; BC025174; AAH25174.1; -.
KW   Hypothetical protein.
SQ   SEQUENCE   243 AA;  25420 MW;  498129CD051DB97B CRC64;

  Query Match           94.6%;  Score 1254;  DB 11;  Length 243;
  Best Local Similarity  93.4%;  Pred. No. 5.1e-100;
  Matches  227;  Conservative    7;  Mismatches    9;  Indels    0;  Gaps  . 0;

Qy          1 MRPLLVLLLLGLAAGSPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPG 60
              ||||| |||||| :|||||||||||||||| |||||||||||||||||||||||||||||
Db          1 MRPLLALLLLGLVSGSPPLDDNKIPSLCPGQPGLPGTPGHHGSQGLPGRDGRDGRDGAPG 60

Qy         61 APGEKGEGGRPGLPGPRGDPGPRGEAGPAGPTGPAGECSVPPRSAFSAKRSESRVPPPSD 120
              |||||||||||||||||:|||||||| |  |||||||||||||||||||||||||:|
Db         61 APGEKGEGGRPGLPGPRGEPGPRGEAGPMGAIGPAGECSVPPRSAFSAKRSESRVPPPAD 120

Qy        121 APLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGESIASFFQ 180
              ||||||||:|||||:|  |||||||||||||||||||||||||||||:|||||||
```

```
Db        121 TPLPFDRVLLNEQGHFDPTTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGQSIASFFQ 180

Qy        181 FFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWHSSP 240
              :||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 YFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWHSSP 240

Qy        241 VFA 243
              |||
Db        241 VFA 243


RESULT 2
Q95MQ4
ID   Q95MQ4      PRELIMINARY;      PRT;   240 AA.
AC   Q95MQ4;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Adipose tissue-specific protein adipo Q.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21369933; PubMed=11382781;
RA   Sato C., Yasukawa Z., Honda N., Matsuda T., Kitajima K.;
RT   "Identification and Adipocyte Differentiation-dependent Expression of
RT   the Unique Disialic Acid Residue in an Adipose Tissue-specific
RT   Glycoprotein, Adipo Q.";
RL   J. Biol. Chem. 276:28849-28856(2001).
DR   EMBL; AF269230; AAK58902.1; -.
DR   InterPro; IPR001073; C1q.
DR   InterPro; IPR000087; Collagen.
DR   Pfam; PF00386; C1q; 1.
DR   Pfam; PF01391; Collagen; 1.
DR   ProDom; PD000007; Collagen; 1.
DR   PROSITE; PS01113; C1Q; UNKNOWN_1.
SQ   SEQUENCE   240 AA;  26091 MW;  C6253BA803B9A668 CRC64;

  Query Match          32.1%;  Score 425.5;  DB 6;  Length 240;
  Best Local Similarity   41.5%;  Pred. No. 7.1e-29;
  Matches  100;  Conservative   35;  Mismatches   81;  Indels   25;  Gaps   10;

Qy          6 VLLLLGLAA-GSPPLDDNKIP-SLCPG-HPGLPGTPGHHGSQGLPGRDGRDGRDGAPGAP 62
              :|||| | : |   ::|  :|   | |   |:|| |||:|:   |||||||      | |
Db          7 LLLLLALPSHGEDNMEDPPLPKGACAGWMAGIPGHPGHNGT---PGRDGRD------GTP 57

Qy         63 GEKGEGGRPGLPGPRGD---------PGPRGEAGPAGPTGPAGECSVPPRSAFSAKRSES 113
              ||||| |  || ||:|:          ||||  |  |  |  || :  |||||    |:
Db         58 GEKGEKGDAGLLGPKGETGDVGMTGAEGPRGFPGTPGRKGEPGEAAYVYRSAFSV-GLET 116

Qy        114 RVPPPSDAPLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGE 173
              ||  | : |: | ::  |:| |||  |||| | :||:|||: | |||   ::  | |  :
Db        117 RVTVP-NVPIRFTKIFYNQQNHYDGSTGKFYCNIPGLYYFSYHITVYMKDVKVSLFKKDK 175

Qy        174 SIASFFQFFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGD-YIGIYASIKTDSTFSGFLV 232
              ::   :  :         ||  :: ||  ||||:||  |: : |:||    ||||:|||:
Db        176 AVLFTYDQYQE-KNVDQASGSVLLHLEVGDQVWLQVYEGENHNGVYADNVNDSTFTGFLL 234

Qy        233 Y 233
              |
Db        235 Y 235


RESULT 3
Q95JD7
ID   Q95JD7      PRELIMINARY;      PRT;   243 AA.
AC   Q95JD7;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Adiponectin.
GN   APM1.
OS   Macaca mulatta (Rhesus macaque).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC   Cercopithecinae; Macaca.
OX   NCBI_TaxID=9544;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=ADIPOSE TISSUE;
RX   MEDLINE=21232234; PubMed=11334417;
RA   Hotta K., Funahashi T., Bodkin N.L., Ortmeyer H.K., Arita Y.,
RA   Hansen B.C., Matsuzawa Y.;
RT   "Circulating concentrations of the adipocyte protein adiponectin are
RT   decreased in parallel with reduced insulin sensitivity during the
RT   progression to type 2 diabetes in rhesus monkeys.";
RL   Diabetes 50:1126-1133(2001).
DR   EMBL; AF404407; AAK92202.1; -.
DR   InterPro; IPR001073; C1q.
DR   InterPro; IPR000087; Collagen.
DR   Pfam; PF00386; C1q; 1.
DR   Pfam; PF01391; Collagen; 1.
DR   ProDom; PD000007; Collagen; 1.
DR   PROSITE; PS01113; C1Q; UNKNOWN_1.
SQ   SEQUENCE   243 AA;  26264 MW;  49A45DAF2B4613FD CRC64;

  Query Match          31.8%;  Score 422;  DB 6;  Length 243;
  Best Local Similarity   42.3%;  Pred. No. 1.4e-28;
  Matches  104;  Conservative   28;  Mismatches   88;  Indels   26;  Gaps    8;

Qy          6 VLLLLGLAAGSPPLDDNKIPSLCPGHPG-----LPGTPGHHGSQGLPGRDGRDGRDGAPG 60
              ||||| | :             | |   |     : | ||| |  |:|||||||      |
Db          6 VLLLLALPSHGQDTTTQGPGVLLPLPKGACTGWMAGIPGHPGHNGVPGRDGRD------G 59

Qy         61 APGEKGEGGRPGLPGPRGD---------PGPRGEAGPAGPTGPAGECSVPPRSAFSAKRS 111
              ||||| | ||| ||:||          ||||  |  |  |  || :  |||||
Db         60 TPGEKGEKGDPGLIGPKGDTGETGVTGAEGPRGFPGIQGRKGEPGEGAYVYRSAFSV-GL 118

Qy        112 ESRVPPPSDAPLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKN 171
              |: |  | : |: | ::  |:| |||  |||| | :||:|||| | |||   ::  | |
Db        119 ETYVTVP-NMPIRFTKIFYNQQNHYDGSTGKFHCNIPGLYYFAYHITVYMKDVKVSLFKK 177

Qy        172 GES-IASFFQFFGGWPKPASLSGGAMVRLEPEDQVWVQV-GVGDYIGIYASIKTDSTFSG 229
              :: : :: |:        |  ||  :: ||  ||||:|| | |:  |:||    ||||:|
Db        178 DKAMLFTYDQYQENNVDQA--SGSVLLHLEVGDQVWLQVYGEGERNGLYADNDNDSTFTG 235

Qy        230 FLVYSD 235
              ||:| |
Db        236 FLLYHD 241


RESULT 4
Q9N178
ID   Q9N178      PRELIMINARY;      PRT;   675 AA.
AC   Q9N178;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Type X collagen.
GN   COL10A1.
OS   Sus scrofa (Pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX   NCBI_TaxID=9823;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21015405; PubMed=11130976;
RA   Nielsen V.H., Bendixen C., Arnbjerg J., Sorensen C.M., Jensen H.E.,
RA   Shukri N.M., Thomsen B.;
RT   "Abnormal growth plate function in pigs carrying a dominant mutation
RT   in type X collagen.";
RL   Mamm. Genome 11:1087-1092(2000).
DR   EMBL; AF222861; AAF37271.1; -.
```

```
DR   InterPro; IPR001073; C1q.
DR   InterPro; IPR000087; Collagen.
DR   Pfam; PF00386; C1q; 1.
DR   Pfam; PF01391; Collagen; 8.
DR   PRINTS; PR00007; COMPLEMNTC1Q.
DR   ProDom; PD000007; Collagen; 2.
DR   SMART; SM00110; C1Q; 1.
DR   PROSITE; PS01113; C1Q; 1.
KW   Collagen.
SQ   SEQUENCE   675 AA;  65447 MW;  26397B10310383F9 CRC64;

  Query Match           31.4%;  Score 415.5;  DB 6;  Length 675;
  Best Local Similarity  36.1%;  Pred. No. 1.7e-27;
  Matches   99;  Conservative   34;  Mismatches   82;  Indels   59;  Gaps    7;

Qy         15 GSPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPGAPGEKGEGGRPGLP 74
              |:| |   |       | |||||  |  |::|:|| :|  |  |||| || :|  | ||:|
Db        402 GNPGLPGPKGDPGIGGPPGLPGPVGPAGAKGVPGHNGEAGPRGAPGIPGTRGPIGPPGIP 461

Qy         75 ----------------------------GPRGDPGPRGEA------GPAGPTGPAGECS 99
                                          || | |||:| |      || || || |: :
Db        462 GFPGSKGDPGNPGPPGPAGIATKGLNGPTGPPGPPGPKGHAGEPGLPGPPGPPGPPGQ-A 520

Qy        100 VPP--------RSAFSAKRSESRVP------------PPSDAPLPFDRVLVNEQGHYDAV 139
              |||        |:  || :  : :|            |   ||:|||::| | | |||
Db        521 VPPEGFVKEGQRAFVSANQGVTGMPVSAFTVILSKAYPAIGAPIPFDKILYNGQQHYDPK 580

Qy        140 TGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGESIA-SFFQFFGGWPKPASLSGGAMVR 198
              || |||::||:|||: |  |         | |||  :  :: ::  |:   |  || |::
Db        581 TGIFTCRIPGIYYFSYHIHVKGTHAWVGLYKNGTPVMYTYDEYVKGYLDQA--SGSAILD 638

Qy        199 LEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLV 232
              |   ||||:|:      |:|:|    |:||||||
Db        639 LTDNDQVWLQLPNAGSNGLYSSEYVHSSFSGFLV 672


RESULT 5
Q9Z1K4
ID   Q9Z1K4      PRELIMINARY;      PRT;   295 AA.
AC   Q9Z1K4;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Collagen alpha 1 type X (Fragment).
GN   COL10A1.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SPRAGUE-DAWLEY;
RX   MEDLINE=20310874; PubMed=10853827;
RA   Marks S.A., Lundmark C., Christersson C., Wurtz T., Odgren P.R.,
RA   Seifert M.F., MacKay C.A., Mason-Savas A., Popoff S.E.;
RT   "Endochondral bone formation in toothless (osteopetrotic) rats:
RT   failures of chondrocyte patterning and type X collagen expression.";
RL   Int. J. Dev. Biol. 44:309-316(2000).
DR   EMBL; AJ131848; CAA10518.1; -.
DR   InterPro; IPR001073; C1q.
DR   InterPro; IPR000087; Collagen.
DR   Pfam; PF00386; C1q; 1.
DR   Pfam; PF01391; Collagen; 3.
DR   PRINTS; PR00007; COMPLEMNTC1Q.
DR   SMART; SM00110; C1Q; 1.
DR   PROSITE; PS01113; C1Q; 1.
FT   NON_TER       1      1
SQ   SEQUENCE   295 AA;  30012 MW;  FF43B1548028813E CRC64;

  Query Match           30.3%;  Score 402;  DB 11;  Length 295;
  Best Local Similarity  36.7%;  Pred. No. 9.3e-27;
  Matches   94;  Conservative   31;  Mismatches   79;  Indels   52;  Gaps    7;

Qy         25 PSLCPGH---------PGLPGTPGHHGSQGLPGRDGRDGRDGAPGAPGE--------KGE 67
              |   |||          ||:||| |  |  |:||  |  |  | |||||          |
Db         41 PKGVPGHNGEAGPRGEPGIPGTRGPIGPPGIPGFPGSKGDPGKPGAPGPAGIVTKGLNGP 100

Qy         68 GGRPGLPGPR---GDPGPRGEAGPAGPTGPAGECSVP---------------------- 101
              | || ||||   |:|| |  || || ||  :  :|
Db        101 AGPPGRPGPRGHTGEPGLPGPPGPPGPPGPPSQAVIPDGFTKSGQRPRLSGMPLVSANQG 160

Qy        102 ----PRSAFSAKRSESRVPPPSDAPLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHA 157
                  | |||:  |::  |  ||:||| :| | | |||  :| |||:||:|||: |
Db        161 VTGMPVSAFTVILSKAY--PAVGAPIPFDEILYNRQQHYDPRSGIFTCKIPGIYYFSYHI 218

Qy        158 TVYRASLQFDLVKNG-ESIASFFQFFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIG 216
              |    :  | |||  :: :: ::  |:  |  || |:: |   ||||:|:   :  |
Db        219 HVKGTHVWVGLYKNGTPTMYTYDEYSKGYLDQA--SGSAIMELTENDQVWLQLPNAESNG 276

Qy        217 IYASIKTDSTFSGFLV 232
              :|:|    |:||||||
Db        277 LYSSEYVHSSFSGFLV 292


RESULT 6
Q9D8U4
ID   Q9D8U4      PRELIMINARY;      PRT;   294 AA.
AC   Q9D8U4;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   1810033K05Rik protein.
GN   1810033K05RIK.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=PANCREAS;
RX   MEDLINE=21085660; PubMed=11217851;
RA   Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA   Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA   Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA   Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA   Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA   Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA   Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA   Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA   Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA   Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA   Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA   Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA   Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA   Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA   Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA   Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA   Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA   Hayashizaki Y.;
RT   "Functional annotation of a full-length mouse cDNA collection.";
RL   Nature 409:685-690(2001).
DR   EMBL; AK007683; BAB25187.1; -.
DR   MGD; MGI:1916433; 1810033K05Rik.
DR   InterPro; IPR001073; C1q.
DR   InterPro; IPR000087; Collagen.
DR   Pfam; PF00386; C1q; 1.
DR   Pfam; PF01391; Collagen; 2.
DR   PRINTS; PR00007; COMPLEMNTC1Q.
DR   SMART; SM00110; C1Q; 1.
DR   PROSITE; PS01113; C1Q; 1.
SQ   SEQUENCE   294 AA;  30865 MW;  6D3905AE7C19E6FA CRC64;

  Query Match           29.8%;  Score 395;  DB 11;  Length 294;
  Best Local Similarity  35.9%;  Pred. No. 3.7e-26;
  Matches   94;  Conservative   31;  Mismatches   93;  Indels   44;  Gaps    6;
```

```
Qy          3 PLLVLLLLGLAAGSPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPGAP 62
              | ||  | |              |   || || ||: |  |· | ||:||:||:|| |
Db         41 PQLVCSLPG------------PQGPPGPPGAPGSSGVVGRMGFPGKDGQDGQDGDRGDS 87

Qy         63 GEKGEGGRP---GLPGPRGDPGPRGEAGPAGP---TGPAGECSVPPRSAFSAKRSESRVP 116
              ||:|  ||    |  ||:|  |  | ||| ||   :|  |:   | :     |: |  :|
Db         88 GEEGPPGRTGNRGKQGPKGKAGAIGRAGPRGPKGVSGTPGKHGTPGKKGPKGKKGEPGLP 147

Qy        117 PPSDA---------------------PLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFA 154
               |                        |: ||::|:|| |||:| :||| | |||:|||
Db        148 GPCSCGSSRAKSAFSVAVTKSYPRERLPIKFDKILMNEGGHYNASSGKFVCSVPGIYYFT 207

Qy        155 VHATVYRASLQFDLVKNGESIASFFQFFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDY 214
                 |:    |   || ||:     |    |    |  ||  :: |:  |:||:|:   :
Db        208 YDITLANKHLAIGLVHNGQYRIRTFDANTGNHDVA--SGSTILALKEGDEVWLQIFYSEQ 265

Qy        215 IGI-YASIKTDSTFSGFLVYSD 235
               |: |    ||| |:|||:|:|
Db        266 NGLFYDPYWTDSLFTGFLIYAD 287


RESULT 7
Q95J95
ID   Q95J95      PRELIMINARY;      PRT;   194 AA.
AC   Q95J95;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Adiponectin (Fragment).
GN   APM1.
OS   Canis familiaris (Dog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX   NCBI_TaxID=9615;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=ADIPOSE TISSUE;
RA   Kabir M., Ananthnarayan S., Ionut V., Kim S.P., Van Citters G.W.,
RA   Dea M.K., Bergman R.N.;
RT   "Regulation of Adiponectin gene expression in the fat-fed dog.";
RL   Submitted (SEP-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF417206; AAL09702.1; -.
DR   InterPro; IPR001073; C1q.
DR   InterPro; IPR000087; Collagen.
DR   Pfam; PF00386; C1q; 1.
DR   Pfam; PF01391; Collagen; 1.
DR   ProDom; PD000007; Collagen; 1.
DR   PROSITE; PS01113; C1Q; UNKNOWN_1.
FT   NON_TER       1      1
FT   NON_TER     194    194
SQ   SEQUENCE   194 AA;  20890 MW;  3AA3D947D187AF9A CRC64;

  Query Match             28.2%;  Score 373.5;  DB 6;  Length 194;
  Best Local Similarity   42.3%;  Pred. No. 1.6e-24;
  Matches   90;  Conservative   22;  Mismatches   66;  Indels   35;  Gaps    9;

Qy          8 LLLGLAAGSPPLDDNKIPSLCPG-HPGLPGTPGHHGSQGLPGRDGRDGRDGAPGAPGEKG 66
              :|| |  |:           |||   |:|| |||:|:   |||||||      | |||||
Db          4 VLLPLPKGA-----------CPGWMAGIPGHPGHNGT---PGRDGRD------GTPGEKG 43

Qy         67 EGGRPGLPGPRGD---------PGPRGEAGPAGPTGPAGECSVPPRSAFSAKRSESRVPP 117
              | | ||| ||:||          ||||  |  |  |  || :   |||||    |||:
Db         44 EKGDPGLVGPKGDTGETGVTGVEGPRGFPGTPGRKGEPGESAYVHRSAFSV-GLESRITV 102

Qy        118 PSDAPLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGESIAS 177
              | : |: | ::   | | |||  |||| | :||:|||: | |||   ::  | |  :  |
Db        103 P-NVPIRFTKIFYNLQNHYDGTTGKFHCNIPGLYYFSYHITVYLKDVKVSLYKKDK--AM 159

Qy        178 FFQFFGGWPKPA-SLSGGAMVRLEPEDQVWVQV 209
                | :     |    ||  :: ||  ||||:||
Db        160 LFTYDQYQEKNVDQASGSVLLHLEVGDQVWLQV 192


RESULT 8
Q9D2V4
ID   Q9D2V4      PRELIMINARY;      PRT;   744 AA.
AC   Q9D2V4;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   Procollagen, type VIII, alpha 1.
GN   COL8A1.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=KIDNEY;
RX   MEDLINE=21085660; PubMed=11217851;
RA   Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA   Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA   Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA   Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA   Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA   Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA   Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA   Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA   Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA   Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA   Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA   Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA   Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA   Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA   Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA   Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA   Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA   Hayashizaki Y.;
RT   "Functional annotation of a full-length mouse cDNA collection.";
RL   Nature 409:685-690(2001).
DR   EMBL; AK018742; BAB31383.1; -.
DR   MGD; MGI:88463; Col8a1.
DR   InterPro; IPR001073; C1q.
DR   InterPro; IPR000087; Collagen.
DR   Pfam; PF00386; C1q; 1.
DR   Pfam; PF01391; Collagen; 7.
DR   PRINTS; PR00007; COMPLEMNTC1Q.
DR   SMART; SM00110; C1Q; 1.
DR   PROSITE; PS01113; C1Q; 1.
SQ   SEQUENCE   744 AA;  73581 MW;  C659BDCCBCD6EB9C CRC64;

  Query Match             28.0%;  Score 370.5;  DB 11;  Length 744;
  Best Local Similarity   34.6%;  Pred. No. 1.4e-23;
  Matches   93;  Conservative   32;  Mismatches   83;  Indels   61;  Gaps    7;

Qy         24 IPSLC--PGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPGAPGEKGEGGRPGLP------- 74
              :| |    | ||:||  |  |  |:||  |  |  | || || ||| | || |
Db        476 VPGLLGPKGEPGIPGDQGLQGPPGIPGIVGPSGPIGPPGIPGPKGEPGLPGPPGFPGVGK 535

Qy         75 ----------------GPRGDPGPRGEAGPAGPTGP-----------------AGECSV 100
                             ||:| ||  |  || || ||                 |  |
Db        536 PGVAGLHGPPGKPGALGPQGQPGLPGPPGPPGPPGPPAVMPTPSPQGEYLPDMGLGIDGV 595

Qy        101 PPRSAFSAKR-------------SESRVP-PPSDAPLPFDRVLVNEQGHYDAVTGKFTC 145
               |  |:: |:                :|  || ||  ||: ||::| | : :|:  || |||
Db        596 KPPHAYAGKKGKHGGPAYEMPAFTAELTVPFPPVGAPVKFDKLLYNGRQNYNPQTGIFTC 655

Qy        146 QVPGVYYFAVHATVYRASLQFDLVKNGESIA-SFFQFFGGWPKPASLSGGAMVRLEPEDQ 204
              :|||||||| |      ::   | || | :  :: ::  |:   |  || |:: | | ||
Db        656 EVPGVYYFAYHVHCKGGNVWVALFKNNEPMMYTYDEYKKGFLDQA--SGSAVLLLRPGDQ 713

Qy        205 VWVQVGVGDYIGIYASIKTDSTFSGFLVY 233
              |::|:      |:||     |:|||:|:|
Db        714 VFLQMPSEQAAGLYAGQYVHSSFSGYLLY 742
```

```
RESULT 9
Q921S8
ID   Q921S8      PRELIMINARY;      PRT;   744 AA.
AC   Q921S8;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Procollagen, type VIII, alpha 1.
GN   COL8A1.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Strausberg R.;
RL   Submitted (JUL-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC011061; AAH11061.1; -.
DR   MGD; MGI:88463; Col8a1.
DR   InterPro; IPR001073; C1q.
DR   InterPro; IPR000087; Collagen.
DR   Pfam; PF00386; C1q; 1.
DR   Pfam; PF01391; Collagen; 7.
DR   PROSITE; PS01113; C1Q; UNKNOWN_1.
KW   Collagen.
SQ   SEQUENCE   744 AA;  73621 MW;  BF7A7FD79D8463AA CRC64;

  Query Match             28.0%;  Score 370.5;  DB 11;  Length 744;
  Best Local Similarity   34.6%;  Pred. No. 1.4e-23;
  Matches   93;  Conservative   32;  Mismatches   83;  Indels   61;  Gaps    7;

Qy         24 IPSLC--PGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPGAPGEKGEGGRPGLP------- 74
              :| |    | ||:||  |  |  |:||  |  |  | || || ||| | || |
Db        476 VPGLLGPKGEPGIPGDQGLQGPPGIPGIVGPSGPIGPPGIPGPKGEPGLPGPPGFPGVGK 535

Qy         75 ---------------GPRGDPGPRGEAGPAGPTGP-----------------AGECSV 100
                             ||:| ||  |  || ||  ||                 |   |
Db        536 PGVAGLHGPPGKPGALGPQGQPGLPGPPGPPGPPGPPAVMPTPSPQGEYLPDMGLGIDGV 595

Qy        101 PPRSAFSAKR-------------SESRVP-PPSDAPLPFDRVLVNEQGHYDAVTGKFTC 145
              |  |:: |:              :|  || ||  ||: ||::| | : :|:. || |||
Db        596 KPPHAYAGKKGKHGGPAYEMPAFTAELTVPFPPVGAPVKFDKLLYNGRQNYNPQTGIFTC 655

Qy        146 QVPGVYYFAVHATVYRASLQFDLVKNGESIA-SFFQFFGGWPKPASLSGGAMVRLEPEDQ 204
              :|||||||| |      ::   | || | :  :: :: |:   |  || |:: | | ||
Db        656 EVPGVYYFAYHVHCKGGNVWVALFKNNEPMMYTYDEYKKGFLDQA--SGSAVLLLRPGDQ 713

Qy        205 VWVQVGVGDYIGIYASIKTDSTFSGFLVY 233
              |::|:      |:||     |:|||:|:|
Db        714 VFLQMPSEQAAGLYAGQYVHSSFSGYLLY 742


RESULT 10
Q96D07
ID   Q96D07      PRELIMINARY;      PRT;   744 AA.
AC   Q96D07;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Hypothetical 73.4 kDa protein.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=LUNG;
RA   Strausberg R.;
RL   Submitted (SEP-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC013581; AAH13581.1; -.
DR   InterPro; IPR001073; C1q.
DR   InterPro; IPR000087; Collagen.
DR   Pfam; PF00386; C1q; 1.
DR   Pfam; PF01391; Collagen; 7.
DR   ProDom; PD000007; Collagen; 2.
DR   PROSITE; PS01113; C1Q; UNKNOWN_1.
KW   Collagen; Hypothetical protein.
SQ   SEQUENCE   744 AA;  73364 MW;  2BC1B0955DE2C9A3 CRC64;

  Query Match             27.8%;  Score 368;  DB 4;  Length 744;
  Best Local Similarity   33.3%;  Pred. No. 2.3e-23;
  Matches   90;  Conservative   34;  Mismatches   84;  Indels   62;  Gaps    6;

Qy         24 IPSLC--PGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPGAPGEKGEGGRPGLP------- 74
              :| |    | ||:||  |  |  |:||  |  |  | || || ||| | || |
Db        475 VPGLLGPKGEPGIPGDQGLQGPPGIPGIGGPSGPIGPPGIPGPKGEPGLPGPPGFPGIGK 534

Qy         75 ---------------GPRGDPGPRGEAGPAGPTGP-----------------AGECS 99
                             ||:| ||  |  || ||  ||                 |
Db        535 PGVAGLHGPPGKPGALGPQGQPGLPGPPGPPGPPGPPAVMPPTPPPQGEYLPDMGLGIDG 594

Qy        100 VPPRSAFSAKRSESRVP---------------PPSDAPLPFDRVLVNEQGHYDAVTGKFT 144
              | |  |: ||: ::  |                ||  ||: |:::| | : :|:  || ||
Db        595 VKPPHAYGAKKGKNGGPAYEMPAFTAELTAPFPPVGAPVKFNKLLYNGRQNYNPQTGIFT 654

Qy        145 CQVPGVYYFAVHATVYRASLQFDLVKNGESIA-SFFQFFGGWPKPASLSGGAMVRLEPED 203
              |:|||||||| |       ::   | || | :  :: :: |:   |  || |:: | | |
Db        655 CEVPGVYYFAYHVHCKGGNVWVALFKNNEPVMYTYDEYKKGFLDQA--SGSAVLLLRPGD 712

Qy        204 QVWVQVGVGDYIGIYASIKTDSTFSGFLVY 233
              :|:::|:      |:||     |:|||:|:|
Db        713 RVFLQMPSEQAAGLYAGQYVHSSFSGYLLY 742


RESULT 11
Q8TEJ5
ID   Q8TEJ5      PRELIMINARY;      PRT;   705 AA.
AC   Q8TEJ5;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   FLJ00201 protein (Fragment).
GN   FLJ00201.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=SPLEEN;
RA   Jikuya H., Takano J., Nomura N., Kikuno R., Nagase T., Ohara O.;
RT   "The nucleotide sequence of a long cDNA clone isolated from human
RT   spleen.";
RL   Submitted (JAN-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AK074129; BAB84955.1; -.
FT   NON_TER       1       1
SQ   SEQUENCE   705 AA;  67430 MW;  07DB85A65A948ED3 CRC64;

  Query Match             27.4%;  Score 363.5;  DB 4;  Length 705;
  Best Local Similarity   39.3%;  Pred. No. 5.2e-23;
  Matches   95;  Conservative   25;  Mismatches   95;  Indels   27;  Gaps    8;

Qy         11 GLAAGSPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPGAPGEKGEGGR 70
              ||   : |:    :| |  | ||||| ||  |  | ||  |  |  | ||:||  |  |
Db        468 GLQGPAGPIGPQGLPGL-KGEPGLPGPPG-EGRAGEPGTAGPTGPPGVPGSPGITGPPGP 525

Qy         71 PGLPGPRGDPGPRGEAGPAG---PTGPA--------------GECSVPPRSAFSAKRSE 112
              || ||| | ||   | | ||   | |                || |     ||:|  :
Db        526 PGPPGPPGAPGAFDETGIAGLHLPNGGVEGAVLGKGGKPQFGLGELSAHATPAFTAVLTS 585

Qy        113 SRVP-PPSDAPLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKN 171
              | | |  |: ||| | |    |:  || ||| | |||||| |  |   ::   |.||
Db        586 ---PFPASGMPVKFDRTLYNGHSGYNPATGIFTCPVGGVYYFAYHVHVKGTNVWVALYKN 642

Qy        172 G-ESIASFFQFFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGF 230
                :  :: ::  |:   |  ||||:::| | ||||||:      |:|::    |:|||
```

```
Db        643 NVPATYTYDEYKKGYLDQA--SGGAVLQLRPNDQVWVQMPSDQANGLYSTEYIHSSFSGF 700

Qy        231 LV 232
              |:
Db        701 LL 702
```

RESULT 12
Q9DCM6
ID  Q9DCM6  PRELIMINARY;  PRT;  245 AA.
AC  Q9DCM6;
DT  01-JUN-2001 (TrEMBLrel. 17, Created)
DT  01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT  01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE  Complement component 1, q subcomponent, alpha polypeptide.
GN  C1QA.
OS  Mus musculus (Mouse).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX  NCBI_TaxID=10090;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=C57BL/6J; TISSUE=KIDNEY;
RX  MEDLINE=21085660; PubMed=11217851;
RA  Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA  Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA  Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA  Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA  Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA  Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA  Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA  Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA  Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA  Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA  Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA  Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA  Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA  Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA  Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA  Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA  Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA  Hayashizaki Y.;
RT  "Functional annotation of a full-length mouse cDNA collection.";
RL  Nature 409:685-690(2001).
RN  [2]
RP  SEQUENCE FROM N.A.
RA  Strausberg R.;
RL  Submitted (JAN-2001) to the EMBL/GenBank/DDBJ databases.
DR  EMBL; AK002655; BAB22262.1; -.
DR  EMBL; BC002086; AAH02086.1; -.
DR  MGD; MGI:88223; C1qa.
DR  InterPro; IPR001073; C1q.
DR  InterPro; IPR000087; Collagen.
DR  Pfam; PF00386; C1q; 1.
DR  Pfam; PF01391; Collagen; 1.
DR  PRINTS; PR00007; COMPLEMNTC1Q.
DR  SMART; SM00110; C1Q; 1.
DR  PROSITE; PS01113; C1Q; 1.
SQ  SEQUENCE  245 AA;  25974 MW;  41C2066D49592020 CRC64;

Query Match  23.7%;  Score 314;  DB 11;  Length 245;
Best Local Similarity  35.3%;  Pred. No. 2.7e-19;
Matches  79;  Conservative  24;  Mismatches  85;  Indels  36;  Gaps  6;

```
Qy         29 PGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPGAPGEKGEGGRPGLPGPRGDPGPRGEAGP 88
              ||:|| || ||  | :| ||  |   | |  |  |: || | || ||  | ||| |  |
Db         36 PGNPGRPGRPGLKGERGEPGAAGI--RTGIRGFKGDPGESGPPGKPGNVGLPGPSGPLGD 93

Qy         89 AGPTGPAGECSVP------PRSAFSAKRSESRVPPPSDAPLPFDRVLVNEQGHYDAVTGK 142
              :|| |  |   |      || |||| |     |      : ||:|| |::  |   ||:
Db         94 SGPQGLKGVKGNPGNIRDQPRPAFSAIRQN---PMTLGNVVIFDKVLTNQESPYQNHTGR 150

Qy        143 FTCQVPGVYYFAVHATVYRASLQFDLVKNGESIASFFQFFGGWPKPA------------ 189
```

```
                | | ||| |||           | ::   :         || |: :
Db        151 FICAVPGFYYF----------NFQVISKWDLCLFIKSSSGGQPRDSLSFSNTNNKGLFQ 199

Qy        190 SLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVY 233
              |:|| :::|   |:||::       ||   : || |||||::
Db .      200 VLAGGTVLQLRRGDEVWIEKDPAKG-RIYQGTEADSIFSGFLIF 242
```

RESULT 13
Q9ES30
ID  Q9ES30  PRELIMINARY;  PRT;  246 AA.
AC  Q9ES30;
DT  01-MAR-2001 (TrEMBLrel. 16, Created)
DT  01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT  01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE  Collagenous repeat-containing sequence of 26kDa protein.
GN  CORS OR CORS26.
OS  Mus musculus (Mouse).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX  NCBI_TaxID=10090;
RN  [1]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=21264842; PubMed=11071891;
RA  Maeda T., Abe M., Kurisu K., Jikko A., Furukawa S.;
RT  "Molecular cloning and characterization of a novel gene, CORS26,
RT  encoding a putative secretory protein and its possible involvement in
RT  skeletal development.";
RL  J. Biol. Chem. 276:3628-3634(2001).
DR  EMBL; AF246265; AAG33704.1; -.
DR  MGD; MGI:1932136; Cors.
DR  InterPro; IPR001073; C1q.
DR  InterPro; IPR000087; Collagen.
DR  Pfam; PF00386; C1q; 1.
DR  Pfam; PF01391; Collagen; 1.
DR  PRINTS; PR00007; COMPLEMNTC1Q.
DR  SMART; SM00110; C1Q; 1.
DR  PROSITE; PS01113; C1Q; 1.
SQ  SEQUENCE  246 AA;  26828 MW;  42A481B3E9F48F7E CRC64;

Query Match  20.4%;  Score 270.5;  DB 11;  Length 246;
Best Local Similarity  31.8%;  Pred. No. 1.5e-15;
Matches  78;  Conservative  33;  Mismatches  107;  Indels  27;  Gaps  6;

```
Qy          4 LLVLLLLGLA----------AGSPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGR 52
              || || |              ||  | | :|      |  |  | ||  |  |:||  |
Db         11 LLPLLFLPFCLCQDEYMESPQAGGLPPDCSKCCHGDYGFRGYQGPPGPPGPPGIPGNHGN 70

Qy         53 DGRDGAPGAPGEKGEGGRPGLPGPRGDPGPRGEAGPAGPTGPAGECSVPPRSAFSAKRSE 112
              :| :|| |  | |||       | :|| ||||| |  || |  |   |||    :  |
Db         71 NGNNGATGHEGAKGE------KGDKGDLGPRGERGQHGPKGEKGYPGVPPELQIAFMASL 124

Qy        113 SRVPPPSDAPLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAV----HATVYRASLQFDL 168
              :      :: : |  |  |    :| :||:|   | |||:|      |  |     :  |
Db        125 ATHFSNQNSGIIFSSVETNIGNFFDVMTGRFGAPVSGVYFFTFSMMKHEDVEEVYVY--L 182

Qy        169 VKNGESIASFFQFFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFS 228
              : || :: | : :       | |  |:::|   |:||:::| |    ::   :  |||:
Db        183 MHNGNTVFSMYSYETKGKSDTS-SNHAVLKLAKGDEVWLRMGNG---ALHGDHQRFSTFA 238

Qy        229 GFLVY 233
              |||::
Db        239 GFLLF 243
```

RESULT 14
Q920N0
ID  Q920N0  PRELIMINARY;  PRT;  196 AA.
AC  Q920N0;
DT  01-DEC-2001 (TrEMBLrel. 19, Created)
DT  01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT  01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE  HP-20.

```
GN   HP-20.
OS   Tamias sibiricus (Siberian chipmunk) (Asian chipmunk).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Sciuridae; Sciurinae;
OC   Tamias.
OX   NCBI_TaxID=64680;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21488336; PubMed=11602349;
RA   Ono M., Hosoe Y., Azuma S., Shoji M., Nara K., Kondo N., Shiba T.,
RA   Takamatsu N.;
RT   "HNF-1 regulates the liver-specific transcription of the chipmunk HP-
RT   20 gene.";
RL   Gene 277:121-127(2001).
DR   EMBL; AB067779; BAB68362.1; -.
DR   InterPro; IPR001073; C1q.
DR   InterPro; IPR000087; Collagen.
DR   Pfam; PF00386; C1q; 1.
DR   Pfam; PF01391; Collagen; 1.
DR   PROSITE; PS01113; C1Q; UNKNOWN_1.
SQ   SEQUENCE   196 AA;  21330 MW;  B07D17EBFA94D2FD CRC64;

  Query Match           20.1%;  Score 266.5;  DB 11;  Length 196;
  Best Local Similarity  37.4%;  Pred. No. 2.5e-15;
  Matches   70;  Conservative   23;  Mismatches   67;  Indels   27;  Gaps    7;

Qy         60 GAPGEKGEGGRPGLPGPRGDPGPRGEA---GPAGPTGPAGECSVPPRSAFSAKRSESRVP 116
              | ||  |  | ||:||||| ||  | |   |  || ||: :|    ||||: | |  |:|
Db         25 GPPGPVGYPGVPGVPGPRGPPGQPGAAGRPGDPGPKGPSVKCPCRERSAFTVKFS-GRLP 83

Qy        117 PPSDAPLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGESIA 176
              |||: |: |  || | |     | || | |  || |:|:    :|   ::  |:||   :
Db         84 PPSE-PVVFTEVLYNTQRDLKASTGVFNCVEPGNYHFSFDVELYHCKVKIGLMKNHIQVM 142

Qy        177 SFFQFFGGWPKPASLSGGAMVR-LEPEDQVWVQVGVGDYIGIYASIKTDST--------F 227
              |      : ||   |||:  |   |:||::          | ::|:          |
Db        143 EKHQLSKNEYENAS---GAMIMPLRQGDKVWLE----------ADVETEEPDQAKVVIYF 189

Qy        228 SGFLVYS 234
              ||||: |
Db        190 SGFLISS 196


RESULT 15
Q9I9Q7
ID   Q9I9Q7      PRELIMINARY;      PRT;   246 AA.
AC   Q9I9Q7;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Mannose binding-like lectin precursor (Fragment).
GN   MBL.
OS   Carassius auratus (Goldfish).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC   Cyprinidae; Carassius.
OX   NCBI_TaxID=7957;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=LIVER;
RX   MEDLINE=20456722; PubMed=11003389;
RA   Vitved L., Holmskov U., Koch C., Teisner B., Hansen S., Skjodt K.;
RT   "The homologue of mannose-binding lectin in the carp family Cyprinidae
RT   is expressed at high level in spleen, and the deduced primary
RT   structure predicts affinity for galactose.";
RL   Immunogenetics 51:955-964(2000).
DR   EMBL; AF227739; AAF63470.1; -.
DR   HSSP; P35247; 1B08.
DR   InterPro; IPR000087; Collagen.
DR   InterPro; IPR001304; Lectin_C.
DR   Pfam; PF01391; Collagen; 2.
DR   Pfam; PF00059; lectin_c; 1.
DR   SMART; SM00034; CLECT; 1.
DR   PROSITE; PS00615; C_TYPE_LECTIN_1; UNKNOWN_1.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
KW   Lectin; Signal.
FT   NON_TER       1      1
FT   SIGNAL       <1     13       POTENTIAL.
FT   VARIANT     145    145       S -> F.
SQ   SEQUENCE   246 AA;  25709 MW;  AB692282D289D0D5 CRC64;

  Query Match           19.1%;  Score 252.5;  DB 13;  Length 246;
  Best Local Similarity  31.9%;  Pred. No. 5.2e-14;
  Matches   80;  Conservative   33;  Mismatches   73;  Indels   65;  Gaps   14;

Qy          4 LLVLLLLGLAAGSPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPGAPG 63
              ||:   | |  |: | : |      || : |:||||||:   ||||||||||:||| |  |
Db          2 LLLQFALQLLDGAEPQNLN-----CPAYGGVPGTPGHN---GLPGRDGRDGKDGAIGPKG 53

Qy         64 EKGEG-----GRPGLPGPRGDPGPRGEAGPAGPTGPAGECSVPPRSAFSAKRSESRVPPP 118
              ||||      | ||  || | | :|| ||:|| |  |  ||        :: ::::
Db         54 EKGESGVSVQGPPGKAGPPGTAGEKGERGPSGPQGSPGSESVLESLKSEIQQLKAKI--- 110

Qy        119 SDAPLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGESIASF 178
                 |::  |:  .|:  |          |  |:     |      ||     : : |
Db        111 ----ATFEK--VSSVCHFRKV---------GQKYYITDGVVG----NFD-----QGLKSC 146

Qy        179 FQFFGGWPKPASLSGGAMV--RLEPEDQVWVQV----GVGD---YIGIYASIKTDSTF-- 227
               :|           || ||  |   |:|  :::    |:|    |||:    ||:  |
Db        147 MEF----------GGTMVSPRTSAENQALLKLVVSSGLGSKKPYIGV-TDRKTEGQFVD 194

Qy        228 --SGFLVYSDW 236
                | :::|
Db        195 TEGKQLTFTNW 205


Search completed: June 18, 2003, 14:52:31
Job time : 83 secs
```

GenCore version 5.1.6


OM protein - protein search, using sw model

Run on: June 18, 2003, 14:53:08 ; Search time 49 Seconds
(without alignments)
536.617 Million cell updates/sec

Title: US-09-943-851A-42
Perfect score: 1325
Sequence: 1 MRPLLVLLLLGLAAGSPPLD.....DSTFSGFLVYSDWHSSPVFA 243

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 417779 seqs, 108206813 residues

Total number of hits satisfying chosen parameters: 417779

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 1325 | 100.0 | 243 | 9 | US-09-944-413-42 | Sequence 42, Appl |
| 2 | 1325 | 100.0 | 243 | 9 | US-09-944-403-42 | Sequence 42, Appl |
| 3 | 1325 | 100.0 | 243 | 9 | US-09-944-896-42 | Sequence 42, Appl |
| 4 | 1325 | 100.0 | 243 | 9 | US-09-944-944-42 | Sequence 42, Appl |
| 5 | 1325 | 100.0 | 243 | 9 | US-09-944-907-42 | Sequence 42, Appl |
| 6 | 1325 | 100.0 | 243 | 9 | US-09-944-929-42 | Sequence 42, Appl |
| 7 | 1325 | 100.0 | 243 | 9 | US-10-028-072-362 | Sequence 362, App |
| 8 | 1325 | 100.0 | 243 | 9 | US-10-121-049-362 | Sequence 362, App |
| 9 | 1325 | 100.0 | 243 | 9 | US-10-123-904-362 | Sequence 362, App |
| 10 | 1325 | 100.0 | 243 | 9 | US-10-140-470-362 | Sequence 362, App |
| 11 | 1325 | 100.0 | 243 | 9 | US-09-796-753-68 | Sequence 68, Appl |
| 12 | 1325 | 100.0 | 243 | 9 | US-10-175-746-362 | Sequence 362, App |
| 13 | 1325 | 100.0 | 243 | 9 | US-10-176-918-362 | Sequence 362, App |
| 14 | 1325 | 100.0 | 243 | 9 | US-10-176-921-362 | Sequence 362, App |
| 15 | 1325 | 100.0 | 243 | 9 | US-10-137-865-362 | Sequence 362, App |
| 16 | 1325 | 100.0 | 243 | 9 | US-10-140-474-362 | Sequence 362, App |
| 17 | 1325 | 100.0 | 243 | 9 | US-10-142-431-362 | Sequence 362, App |
| 18 | 1325 | 100.0 | 243 | 9 | US-10-143-114-362 | Sequence 362, App |
| 19 | 1325 | 100.0 | 243 | 9 | US-10-140-002-362 | Sequence 362, App |
| 20 | 1325 | 100.0 | 243 | 9 | US-10-142-419-362 | Sequence 362, App |
| 21 | 1325 | 100.0 | 243 | 9 | US-10-123-262-362 | Sequence 362, App |
| 22 | 1325 | 100.0 | 243 | 9 | US-10-142-423-362 | Sequence 362, App |
| 23 | 1325 | 100.0 | 243 | 9 | US-10-121-050-362 | Sequence 362, App |
| 24 | 1325 | 100.0 | 243 | 9 | US-10-141-755-362 | Sequence 362, App |
| 25 | 1325 | 100.0 | 243 | 9 | US-10-143-032-362 | Sequence 362, App |
| 26 | 1325 | 100.0 | 243 | 9 | US-10-123-108-362 | Sequence 362, App |
| 27 | 1325 | 100.0 | 243 | 9 | US-10-123-236-362 | Sequence 362, App |
| 28 | 1325 | 100.0 | 243 | 9 | US-10-123-261-362 | Sequence 362, App |
| 29 | 1325 | 100.0 | 243 | 9 | US-10-140-921-362 | Sequence 362, App |
| 30 | 1325 | 100.0 | 243 | 9 | US-10-140-928-362 | Sequence 362, App |
| 31 | 1325 | 100.0 | 243 | 9 | US-10-121-045-362 | Sequence 362, App |
| 32 | 1325 | 100.0 | 243 | 9 | US-10-123-292-362 | Sequence 362, App |
| 33 | 1325 | 100.0 | 243 | 9 | US-10-123-903-362 | Sequence 362, App |
| 34 | 1325 | 100.0 | 243 | 9 | US-10-124-819-362 | Sequence 362, App |
| 35 | 1325 | 100.0 | 243 | 9 | US-10-124-822-362 | Sequence 362, App |
| 36 | 1325 | 100.0 | 243 | 9 | US-10-140-925-362 | Sequence 362, App |
| 37 | 1325 | 100.0 | 243 | 9 | US-10-160-498-362 | Sequence 362, App |
| 38 | 1325 | 100.0 | 243 | 9 | US-09-944-884-42 | Sequence 42, Appl |
| 39 | 1325 | 100.0 | 243 | 9 | US-10-121-041-362 | Sequence 362, App |
| 40 | 1325 | 100.0 | 243 | 9 | US-10-121-043-362 | Sequence 362, App |
| 41 | 1325 | 100.0 | 243 | 9 | US-10-121-047-362 | Sequence 362, App |
| 42 | 1325 | 100.0 | 243 | 9 | US-10-123-215-362 | Sequence 362, App |
| 43 | 1325 | 100.0 | 243 | 9 | US-10-123-902-362 | Sequence 362, App |
| 44 | 1325 | 100.0 | 243 | 9 | US-10-123-908-362 | Sequence 362, App |
| 45 | 1325 | 100.0 | 243 | 9 | US-10-123-909-362 | Sequence 362, App |

ALIGNMENTS

```
RESULT 1
US-09-944-413-42
; Sequence 42, Application US/09944413
; Patent No. US20020156004A1
; GENERAL INFORMATION:
;  APPLICANT: Baker, Kevin
;  APPLICANT:  Botstein,David
;  APPLICANT:  Eaton,Dan
;  APPLICANT:  Ferrara,Napoleone
;  APPLICANT:  Filvaroff,Ellen
;  APPLICANT:  Gerritsen,Mary
;  APPLICANT:  Goddard,Audrey
;  APPLICANT:  Godowski,Paul
;  APPLICANT:  Grimaldi,Christopher
;  APPLICANT:  Gurney,Austin
;  APPLICANT:  Hillan,Kenneth
;  APPLICANT:  Kljavin,Ivar
;  APPLICANT:  Napier,Mary
;  APPLICANT:  Roy,Margaret
;  APPLICANT:  Tumas,Daniel
;  APPLICANT:  Wood,William
;  TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;  TITLE OF INVENTION:  ACIDS ENCODING THE SAME
;  FILE REFERENCE: P2548P1C1
;  CURRENT APPLICATION NUMBER: US/09/944,413
;  CURRENT FILING DATE:  2001-09-26
;  PRIOR APPLICATION NUMBER: 09/866,028
;  PRIOR FILING DATE: 2001-05-25
;  PRIOR APPLICATION NUMBER: 60/067,411
;  PRIOR FILING DATE: December 3, 1997
;  PRIOR APPLICATION NUMBER: 60/069,334
;  PRIOR FILING DATE: December 11, 1997
;  PRIOR APPLICATION NUMBER: 60/069335
;  PRIOR FILING DATE: December 11, 1997
;  PRIOR APPLICATION NUMBER: 60/069,278
;  PRIOR FILING DATE: December 11, 1997
;  PRIOR APPLICATION NUMBER: 60/069,425
;  PRIOR FILING DATE: December 12, 1997
;  PRIOR APPLICATION NUMBER: 60/069,696
;  PRIOR FILING DATE: December 16, 1997
;  PRIOR APPLICATION NUMBER: 60/069,694
;  PRIOR FILING DATE: December 16, 1997
```

; PRIOR APPLICATION NUMBER: 60/069,702
; PRIOR FILING DATE: December 16, 1997
; PRIOR APPLICATION NUMBER: 60/069,870
; PRIOR FILING DATE: December 17, 1997
; PRIOR APPLICATION NUMBER: 60/069,873
; PRIOR FILING DATE: December 17, 1997
; PRIOR APPLICATION NUMBER: 60/068,017
; PRIOR FILING DATE: December 18, 1997
; PRIOR APPLICATION NUMBER: 60/070,440
; PRIOR FILING DATE: January 5, 1998
; PRIOR APPLICATION NUMBER: 60/074,086
; PRIOR FILING DATE: February 9, 1998
; PRIOR APPLICATION NUMBER: 60/074,092
; PRIOR FILING DATE: February 9, 1998
; PRIOR APPLICATION NUMBER: 60/075,945
; PRIOR FILING DATE: February 25, 1998
; PRIOR APPLICATION NUMBER: 60/112,850
; PRIOR FILING DATE: December 16, 1998
; PRIOR APPLICATION NUMBER: 60/113,296
; PRIOR FILING DATE: December 22, 1998
; PRIOR APPLICATION NUMBER: 60/146,222
; PRIOR FILING DATE: July 28, 1999
; PRIOR APPLICATION NUMBER: PCT/US98/19330
; PRIOR FILING DATE: September 16, 1998
; PRIOR APPLICATION NUMBER: PCT/US98/25108
; PRIOR FILING DATE: December 1, 1998
; PRIOR APPLICATION NUMBER: 09/216,021
; PRIOR FILING DATE: December 16, 1998
; PRIOR APPLICATION NUMBER: 09/218,517
; PRIOR FILING DATE: December 22, 1998
; PRIOR APPLICATION NUMBER: 09/254,311
; PRIOR FILING DATE: March 3, 1999
; PRIOR APPLICATION NUMBER: PCT/US99/12252
; PRIOR FILING DATE: June 22, 1999
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: September 15, 1999
; PRIOR APPLICATION NUMBER: PCT/US99/28409
; PRIOR FILING DATE: No. US20020156004Alember 30, 1999
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: No. US20020156004Alember 30, 1999
; PRIOR APPLICATION NUMBER: PCT/US99/28301
; PRIOR FILING DATE: December1, 1999
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: December 16, 1999
; PRIOR APPLICATION NUMBER: PCT/US00/03565
; PRIOR FILING DATE: February 11, 2000
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: February 22, 2000
; PRIOR APPLICATION NUMBER: PCT/US00/05841
; PRIOR FILING DATE: March 2, 2000
; PRIOR APPLICATION NUMBER: PCT/US00/08439
; PRIOR FILING DATE: March 30, 2000
; PRIOR APPLICATION NUMBER: PCT/US00/14042
; PRIOR FILING DATE: May 22, 2000
; PRIOR APPLICATION NUMBER: PCT/US00/20710
; PRIOR FILING DATE: July 28, 2000
; PRIOR APPLICATION NUMBER: PCT/US00/32678
; PRIOR FILING DATE: December 1, 2000
; PRIOR APPLICATION NUMBER: PCT/US01/06520
; PRIOR FILING DATE: February 28, 2001
; NUMBER OF SEQ ID NOS: 120
; SEQ ID NO 42
; LENGTH: 243
; TYPE: PRT
; ORGANISM: Homo Sapien
US-09-944-413-42

Query Match 100.0%; Score 1325; DB 9; Length 243;
Best Local Similarity 100.0%; Pred. No. 2.3e-89;
Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MRPLLVLLLLGLAAGSPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MRPLLVLLLLGLAAGSPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPG 60

Qy         61 APGEKGEGGRPGLPGPRGDPGPRGEAGPAGPTGPAGECSVPPRSAFSAKRSESRVPPPSD 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 APGEKGEGGRPGLPGPRGDPGPRGEAGPAGPTGPAGECSVPPRSAFSAKRSESRVPPPSD 120

Qy        121 APLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGESIASFFQ 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 APLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGESIASFFQ 180

Qy        181 FFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWHSSP 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 FFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWHSSP 240

Qy        241 VFA 243
              |||
Db        241 VFA 243
```

RESULT 2
US-09-944-403-42
; Sequence 42, Application US/09944403
; Patent No. US20020165143A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin
; APPLICANT: Botstein,David
; APPLICANT: Eaton,Dan
; APPLICANT: Ferrara,Napoleone
; APPLICANT: Filvaroff,Ellen
; APPLICANT: Gerritsen,Mary
; APPLICANT: Goddard,Audrey
; APPLICANT: Godowski,Paul
; APPLICANT: Grimaldi,Christopher
; APPLICANT: Gurney,Austin
; APPLICANT: Hillan,Kenneth
; APPLICANT: Kljavin,Ivar
; APPLICANT: Napier,Mary
; APPLICANT: Roy,Margaret
; APPLICANT: Tumas,Daniel
; APPLICANT: Wood,William
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P2548P1C1
; CURRENT APPLICATION NUMBER: US/09/944,403
; CURRENT FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 09/866,028
; PRIOR FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: 60/067,411
; PRIOR FILING DATE: December 3, 1997
; PRIOR APPLICATION NUMBER: 60/069,334
; PRIOR FILING DATE: December 11, 1997
; PRIOR APPLICATION NUMBER: 60/069335
; PRIOR FILING DATE: December 11, 1997
; PRIOR APPLICATION NUMBER: 60/069,278
; PRIOR FILING DATE: December 11, 1997
; PRIOR APPLICATION NUMBER: 60/069,425
; PRIOR FILING DATE: December 12, 1997
; PRIOR APPLICATION NUMBER: 60/069,696
; PRIOR FILING DATE: December 16, 1997
; PRIOR APPLICATION NUMBER: 60/069,694
; PRIOR FILING DATE: December 16, 1997
; PRIOR APPLICATION NUMBER: 60/069,702
; PRIOR FILING DATE: December 16, 1997
; PRIOR APPLICATION NUMBER: 60/069,870
; PRIOR FILING DATE: December 17, 1997
; PRIOR APPLICATION NUMBER: 60/069,873
; PRIOR FILING DATE: December 17, 1997
; PRIOR APPLICATION NUMBER: 60/068,017
; PRIOR FILING DATE: December 18, 1997
; PRIOR APPLICATION NUMBER: 60/070,440
; PRIOR FILING DATE: January 5, 1998
; PRIOR APPLICATION NUMBER: 60/074,086
; PRIOR FILING DATE: February 9, 1998

```
;  PRIOR APPLICATION NUMBER: 60/074,092
;  PRIOR FILING DATE: February 9, 1998
;  PRIOR APPLICATION NUMBER: 60/075,945
;  PRIOR FILING DATE: February 25, 1998
;  PRIOR APPLICATION NUMBER: 60/112,850
;  PRIOR FILING DATE: December 16, 1998
;  PRIOR APPLICATION NUMBER: 60/113,296
;  PRIOR FILING DATE: December 22, 1998
;  PRIOR APPLICATION NUMBER: 60/146,222
;  PRIOR FILING DATE: July 28, 1999
;  PRIOR APPLICATION NUMBER: PCT/US98/19330
;  PRIOR FILING DATE: September 16, 1998
;  PRIOR APPLICATION NUMBER: PCT/US98/25108
;  PRIOR FILING DATE: December 1, 1998
;  PRIOR APPLICATION NUMBER: 09/216,021
;  PRIOR FILING DATE: December 16, 1998
;  PRIOR APPLICATION NUMBER: 09/218,517
;  PRIOR FILING DATE: December 22, 1998
;  PRIOR APPLICATION NUMBER: 09/254,311
;  PRIOR FILING DATE: March 3, 1999
;  PRIOR APPLICATION NUMBER: PCT/US99/12252
;  PRIOR FILING DATE: June 22, 1999
;  PRIOR APPLICATION NUMBER: PCT/US99/21090
;  PRIOR FILING DATE: September 15, 1999
;  PRIOR APPLICATION NUMBER: PCT/US99/28409
;  PRIOR FILING DATE: No. US20020165143A1ember 30, 1999
;  PRIOR APPLICATION NUMBER: PCT/US99/28313
;  PRIOR FILING DATE: No. US20020165143A1ember 30, 1999
;  PRIOR APPLICATION NUMBER: PCT/US99/28301
;  PRIOR FILING DATE: December1, 1999
;  PRIOR APPLICATION NUMBER: PCT/US99/30095
;  PRIOR FILING DATE: December 16, 1999
;  PRIOR APPLICATION NUMBER: PCT/US00/03565
;  PRIOR FILING DATE: February 11, 2000
;  PRIOR APPLICATION NUMBER: PCT/US00/04414
;  PRIOR FILING DATE: February 22, 2000
;  PRIOR APPLICATION NUMBER: PCT/US00/05841
;  PRIOR FILING DATE: March 2, 2000
;  PRIOR APPLICATION NUMBER: PCT/US00/08439
;  PRIOR FILING DATE: March 30, 2000
;  PRIOR APPLICATION NUMBER: PCT/US00/14042
;  PRIOR FILING DATE: May 22, 2000
;  PRIOR APPLICATION NUMBER: PCT/US00/20710
;  PRIOR FILING DATE: July 28, 2000
;  PRIOR APPLICATION NUMBER: PCT/US00/32678
;  PRIOR FILING DATE: December 1, 2000
;  PRIOR APPLICATION NUMBER: PCT/US01/06520
;  PRIOR FILING DATE: February 28, 2001
;  NUMBER OF SEQ ID NOS: 120
; SEQ ID NO 42
;   LENGTH: 243
;   TYPE: PRT
;   ORGANISM: Homo Sapien
US-09-944-403-42

  Query Match             100.0%;  Score 1325;  DB 9;  Length 243;
  Best Local Similarity   100.0%;  Pred. No. 2.3e-89;
  Matches  243;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MRPLLVLLLLGLAAGSPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MRPLLVLLLLGLAAGSPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPG 60

Qy         61 APGEKGEGGRPGLPGPRGDPGPRGEAGPAGPTGPAGECSVPPRSAFSAKRSESRVPPPSD 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 APGEKGEGGRPGLPGPRGDPGPRGEAGPAGPTGPAGECSVPPRSAFSAKRSESRVPPPSD 120

Qy        121 APLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGESIASFFQ 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 APLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGESIASFFQ 180

Qy        181 FFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWHSSP 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 FFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWHSSP 240

Qy        241 VFA 243
              |||
Db        241 VFA 243


RESULT 3
US-09-944-896-42
; Sequence 42, Application US/09944896
; Patent No. US20020168715A1
; GENERAL INFORMATION:
;  APPLICANT: Baker, Kevin
;  APPLICANT:  Botstein,David
;  APPLICANT:  Eaton,Dan
;  APPLICANT:  Ferrara,Napoleone
;  APPLICANT:  Filvaroff,Ellen
;  APPLICANT:  Gerritsen,Mary
;  APPLICANT:  Goddard,Audrey
;  APPLICANT:  Godowski,Paul
;  APPLICANT:  Grimaldi,Christopher
;  APPLICANT:  Gurney,Austin
;  APPLICANT:  Hillan,Kenneth
;  APPLICANT:  Kljavin,Ivar
;  APPLICANT:  Napier,Mary
;  APPLICANT:  Roy,Margaret
;  APPLICANT:  Tumas,Daniel
;  APPLICANT:  Wood,William
;  TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;  TITLE OF INVENTION:  ACIDS ENCODING THE SAME
;  FILE REFERENCE: P2548P1C1
;  CURRENT APPLICATION NUMBER: US/09/944,896
;  CURRENT FILING DATE:  2001-08-31
;  PRIOR APPLICATION NUMBER: 09/866,028
;  PRIOR FILING DATE: 2001-05-25
;  PRIOR APPLICATION NUMBER: 60/069,334
;  PRIOR FILING DATE: December 11, 1997
;  PRIOR APPLICATION NUMBER: 60/069335
;  PRIOR FILING DATE: December 11, 1997
;  PRIOR APPLICATION NUMBER: 60/069,278
;  PRIOR FILING DATE: December 11, 1997
;  PRIOR APPLICATION NUMBER: 60/069,425
;  PRIOR FILING DATE: December 12, 1997
;  PRIOR APPLICATION NUMBER: 60/069,696
;  PRIOR FILING DATE: December 16, 1997
;  PRIOR APPLICATION NUMBER: 60/069,694
;  PRIOR FILING DATE: December 16, 1997
;  PRIOR APPLICATION NUMBER: 60/069,702
;  PRIOR FILING DATE: December 16, 1997
;  PRIOR APPLICATION NUMBER: 60/069,870
;  PRIOR FILING DATE: December 17, 1997
;  PRIOR APPLICATION NUMBER: 60/069,873
;  PRIOR FILING DATE: December 17, 1997
;  PRIOR APPLICATION NUMBER: 60/068,017
;  PRIOR FILING DATE: December 18, 1997
;  PRIOR APPLICATION NUMBER: 60/070,440
;  PRIOR FILING DATE: January 5, 1998
;  PRIOR APPLICATION NUMBER: 60/074,086
;  PRIOR FILING DATE: February 9, 1998
;  PRIOR APPLICATION NUMBER: 60/074,092
;  PRIOR FILING DATE: February 9, 1998
;  PRIOR APPLICATION NUMBER: 60/075,945
;  PRIOR FILING DATE: February 25, 1998
;  PRIOR APPLICATION NUMBER: 60/112,850
;  PRIOR FILING DATE: December 16, 1998
;  PRIOR APPLICATION NUMBER: 60/113,296
;  PRIOR FILING DATE: December 22, 1998
;  PRIOR APPLICATION NUMBER: 60/146,222
;  PRIOR FILING DATE: July 28, 1999
;  PRIOR APPLICATION NUMBER: PCT/US98/19330
;  PRIOR FILING DATE: September 16, 1998
;  PRIOR APPLICATION NUMBER: PCT/US98/25108
;  PRIOR FILING DATE: December 1, 1998
```

```
;  PRIOR APPLICATION NUMBER: 09/216,021
;  PRIOR FILING DATE: December 16, 1998
;  PRIOR APPLICATION NUMBER: 09/218,517
;  PRIOR FILING DATE: December 22, 1998
;  PRIOR APPLICATION NUMBER: 09/254,311
;  PRIOR FILING DATE: March 3, 1999
;  PRIOR APPLICATION NUMBER: PCT/US99/12252
;  PRIOR FILING DATE: June 22, 1999
;  PRIOR APPLICATION NUMBER: PCT/US99/21090
;  PRIOR FILING DATE: September 15, 1999
;  PRIOR APPLICATION NUMBER: PCT/US99/28409
;  PRIOR FILING DATE: No. US20020168715A1ember 30, 1999
;  PRIOR APPLICATION NUMBER: PCT/US99/28313
;  PRIOR FILING DATE: No. US20020168715A1ember 30, 1999
;  PRIOR APPLICATION NUMBER: PCT/US99/28301
;  PRIOR FILING DATE: December1, 1999
;  PRIOR APPLICATION NUMBER: PCT/US99/30095
;  PRIOR FILING DATE: December 16, 1999
;  PRIOR APPLICATION NUMBER: PCT/US00/03565
;  PRIOR FILING DATE: February 11, 2000
;  PRIOR APPLICATION NUMBER: PCT/US00/04414
;  PRIOR FILING DATE: February 22, 2000
;  PRIOR APPLICATION NUMBER: PCT/US00/05841
;  PRIOR FILING DATE: March 2, 2000
;  PRIOR APPLICATION NUMBER: PCT/US00/08439
;  PRIOR FILING DATE: March 30, 2000
;  PRIOR APPLICATION NUMBER: PCT/US00/14042
;  PRIOR FILING DATE: May 22, 2000
;  PRIOR APPLICATION NUMBER: PCT/US00/20710
;  PRIOR FILING DATE: July 28, 2000
;  PRIOR APPLICATION NUMBER: PCT/US00/32678
;  PRIOR FILING DATE: December 1, 2000
;  PRIOR APPLICATION NUMBER: PCT/US01/06520
;  PRIOR FILING DATE: February 28, 2001
;  NUMBER OF SEQ ID NOS: 120
; SEQ ID NO 42
;   LENGTH: 243
;   TYPE: PRT
;   ORGANISM: Homo Sapien
US-09-944-896-42

  Query Match           100.0%;  Score 1325;  DB 9;  Length 243;
  Best Local Similarity  100.0%;  Pred. No. 2.3e-89;
  Matches  243;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MRPLLVLLLLGLAAGSPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MRPLLVLLLLGLAAGSPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPG 60

Qy         61 APGEKGEGGRPGLPGPRGDPGPRGEAGPAGPTGPAGECSVPPRSAFSAKRSESRVPPPSD 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 APGEKGEGGRPGLPGPRGDPGPRGEAGPAGPTGPAGECSVPPRSAFSAKRSESRVPPPSD 120

Qy        121 APLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGESIASFFQ 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 APLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGESIASFFQ 180

Qy        181 FFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWHSSP 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 FFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWHSSP 240

Qy        241 VFA 243
              |||
Db        241 VFA 243


RESULT 4
US-09-944-944-42
; Sequence 42, Application US/09944944
; Patent No. US20020173463A1
; GENERAL INFORMATION:
;  APPLICANT: Baker, Kevin
;  APPLICANT:  Botstein,David
;  APPLICANT:  Eaton,Dan
;  APPLICANT:  Ferrara,Napoleone
;  APPLICANT:  Filvaroff,Ellen
;  APPLICANT:  Gerritsen,Mary
;  APPLICANT:  Goddard,Audrey
;  APPLICANT:  Godowski,Paul
;  APPLICANT:  Grimaldi,Christopher
;  APPLICANT:  Gurney,Austin
;  APPLICANT:  Hillan,Kenneth
;  APPLICANT:  Kljavin,Ivar
;  APPLICANT:  Napier,Mary
;  APPLICANT:  Roy,Margaret
;  APPLICANT:  Tumas,Daniel
;  APPLICANT:  Wood,William
;  TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;  TITLE OF INVENTION:  ACIDS ENCODING THE SAME
;  FILE REFERENCE: P2548P1C1
;  CURRENT APPLICATION NUMBER: US/09/944,944
;  CURRENT FILING DATE:  2001-09-26
;  PRIOR APPLICATION NUMBER: 09/866,028
;  PRIOR FILING DATE: 2001-05-25
;  PRIOR APPLICATION NUMBER: 60/067,411
;  PRIOR FILING DATE: December 3, 1997
;  PRIOR APPLICATION NUMBER: 60/069,334
;  PRIOR FILING DATE: December 11, 1997
;  PRIOR APPLICATION NUMBER: 60/069335
;  PRIOR FILING DATE: December 11, 1997
;  PRIOR APPLICATION NUMBER: 60/069,278
;  PRIOR FILING DATE: December 11, 1997
;  PRIOR APPLICATION NUMBER: 60/069,425
;  PRIOR FILING DATE: December 12, 1997
;  PRIOR APPLICATION NUMBER: 60/069,696
;  PRIOR FILING DATE: December 16, 1997
;  PRIOR APPLICATION NUMBER: 60/069,694
;  PRIOR FILING DATE: December 16, 1997
;  PRIOR APPLICATION NUMBER: 60/069,702
;  PRIOR FILING DATE: December 16, 1997
;  PRIOR APPLICATION NUMBER: 60/069,870
;  PRIOR FILING DATE: December 17, 1997
;  PRIOR APPLICATION NUMBER: 60/069,873
;  PRIOR FILING DATE: December 17, 1997
;  PRIOR APPLICATION NUMBER: 60/068,017
;  PRIOR FILING DATE: December 18, 1997
;  PRIOR APPLICATION NUMBER: 60/070,440
;  PRIOR FILING DATE: January 5, 1998
;  PRIOR APPLICATION NUMBER: 60/074,086
;  PRIOR FILING DATE: February 9, 1998
;  PRIOR APPLICATION NUMBER: 60/074,092
;  PRIOR FILING DATE: February 9, 1998
;  PRIOR APPLICATION NUMBER: 60/075,945
;  PRIOR FILING DATE: February 25, 1998
;  PRIOR APPLICATION NUMBER: 60/112,850
;  PRIOR FILING DATE: December 16, 1998
;  PRIOR APPLICATION NUMBER: 60/113,296
;  PRIOR FILING DATE: December 22, 1998
;  PRIOR APPLICATION NUMBER: 60/146,222
;  PRIOR FILING DATE: July 28, 1999
;  PRIOR APPLICATION NUMBER: PCT/US98/19330
;  PRIOR FILING DATE: September 16, 1998
;  PRIOR APPLICATION NUMBER: PCT/US98/25108
;  PRIOR FILING DATE: December 1, 1998
;  PRIOR APPLICATION NUMBER: 09/216,021
;  PRIOR FILING DATE: December 16, 1998
;  PRIOR APPLICATION NUMBER: 09/218,517
;  PRIOR FILING DATE: December 22, 1998
;  PRIOR APPLICATION NUMBER: 09/254,311
;  PRIOR FILING DATE: March 3, 1999
;  PRIOR APPLICATION NUMBER: PCT/US99/12252
;  PRIOR FILING DATE: June 22, 1999
;  PRIOR APPLICATION NUMBER: PCT/US99/21090
;  PRIOR FILING DATE: September 15, 1999
;  PRIOR APPLICATION NUMBER: PCT/US99/28409
;  PRIOR FILING DATE: No. US20020173463A1ember 30, 1999
```

```
;  PRIOR APPLICATION NUMBER: PCT/US99/28313
;  PRIOR FILING DATE: No. US20020173463Alember 30, 1999
;  PRIOR APPLICATION NUMBER: PCT/US99/28301
;  PRIOR FILING DATE: Decemberl, 1999
;  PRIOR APPLICATION NUMBER: PCT/US99/30095
;  PRIOR FILING DATE: December 16, 1999
;  PRIOR APPLICATION NUMBER: PCT/US00/03565
;  PRIOR FILING DATE: February 11, 2000
;  PRIOR APPLICATION NUMBER: PCT/US00/04414
;  PRIOR FILING DATE: February 22, 2000
;  PRIOR APPLICATION NUMBER: PCT/US00/05841
;  PRIOR FILING DATE: March 2, 2000
;  PRIOR APPLICATION NUMBER: PCT/US00/08439
;  PRIOR FILING DATE: March 30, 2000
;  PRIOR APPLICATION NUMBER: PCT/US00/14042
;  PRIOR FILING DATE: May 22, 2000
;  PRIOR APPLICATION NUMBER: PCT/US00/20710
;  PRIOR FILING DATE: July 28, 2000
;  PRIOR APPLICATION NUMBER: PCT/US00/32678
;  PRIOR FILING DATE: December 1, 2000
;  PRIOR APPLICATION NUMBER: PCT/US01/06520
;  PRIOR FILING DATE: February 28, 2001
;  NUMBER OF SEQ ID NOS: 120
; SEQ ID NO 42
;   LENGTH: 243
;   TYPE: PRT
;   ORGANISM: Homo Sapien
US-09-944-944-42

  Query Match           100.0%;  Score 1325;  DB 9;  Length 243;
  Best Local Similarity  100.0%;  Pred. No. 2.3e-89;
  Matches  243;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MRPLLVLLLLGLAAGSPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MRPLLVLLLLGLAAGSPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPG 60

Qy         61 APGEKGEGGRPGLPGPRGDPGPRGEAGPAGPTGPAGECSVPPRSAFSAKRSESRVPPPSD 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 APGEKGEGGRPGLPGPRGDPGPRGEAGPAGPTGPAGECSVPPRSAFSAKRSESRVPPPSD 120

Qy        121 APLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGESIASFFQ 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 APLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGESIASFFQ 180

Qy        181 FFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWHSSP 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 FFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWHSSP 240

Qy        241 VFA 243
              |||
Db        241 VFA 243


RESULT 5
US-09-944-907-42
; Sequence 42, Application US/09944907
; Publication No. US20020198147A1
; GENERAL INFORMATION:
;  APPLICANT: Baker, Kevin
;  APPLICANT:  Botstein,David
;  APPLICANT:  Eaton,Dan
;  APPLICANT:  Ferrara,Napoleone
;  APPLICANT:  Filvaroff,Ellen
;  APPLICANT:  Gerritsen,Mary
;  APPLICANT:  Goddard,Audrey
;  APPLICANT:  Godowski,Paul
;  APPLICANT:  Grimaldi,Christopher
;  APPLICANT:  Gurney,Austin
;  APPLICANT:  Hillan,Kenneth
;  APPLICANT:  Kljavin,Ivar
;  APPLICANT:  Napier,Mary
;  APPLICANT:  Roy,Margaret
;  APPLICANT:  Tumas,Daniel
;  APPLICANT:  Wood,William
;  TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;  TITLE OF INVENTION:  ACIDS ENCODING THE SAME
;  FILE REFERENCE: P2548P1C1
;  CURRENT APPLICATION NUMBER: US/09/944,907
;  CURRENT FILING DATE:  2001-08-31
;  PRIOR APPLICATION NUMBER: 09/866,028
;  PRIOR FILING DATE: 2001-05-25
;  NUMBER OF SEQ ID NOS: 120
; SEQ ID NO 42
;   LENGTH: 243
;   TYPE: PRT
;   ORGANISM: Homo Sapien
US-09-944-907-42

  Query Match           100.0%;  Score 1325;  DB 9;  Length 243;
  Best Local Similarity  100.0%;  Pred. No. 2.3e-89;
  Matches  243;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MRPLLVLLLLGLAAGSPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MRPLLVLLLLGLAAGSPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPG 60

Qy         61 APGEKGEGGRPGLPGPRGDPGPRGEAGPAGPTGPAGECSVPPRSAFSAKRSESRVPPPSD 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 APGEKGEGGRPGLPGPRGDPGPRGEAGPAGPTGPAGECSVPPRSAFSAKRSESRVPPPSD 120

Qy        121 APLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGESIASFFQ 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 APLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGESIASFFQ 180

Qy        181 FFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWHSSP 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 FFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWHSSP 240

Qy        241 VFA 243
              |||
Db        241 VFA 243


RESULT 6
US-09-944-929-42
; Sequence 42, Application US/09944929
; Publication No. US20020197612A1
; GENERAL INFORMATION:
;  APPLICANT: Baker, Kevin
;  APPLICANT:  Botstein,David
;  APPLICANT:  Eaton,Dan
;  APPLICANT:  Ferrara,Napoleone
;  APPLICANT:  Filvaroff,Ellen
;  APPLICANT:  Gerritsen,Mary
;  APPLICANT:  Goddard,Audrey
;  APPLICANT:  Godowski,Paul
;  APPLICANT:  Grimaldi,Christopher
;  APPLICANT:  Gurney,Austin
;  APPLICANT:  Hillan,Kenneth
;  APPLICANT:  Kljavin,Ivar
;  APPLICANT:  Napier,Mary
;  APPLICANT:  Roy,Margaret
;  APPLICANT:  Tumas,Daniel
;  APPLICANT:  Wood,William
;  TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;  TITLE OF INVENTION:  ACIDS ENCODING THE SAME
;  FILE REFERENCE: P2548P1C1
;  CURRENT APPLICATION NUMBER: US/09/944,929
;  CURRENT FILING DATE:  2001-08-31
;  PRIOR APPLICATION NUMBER: 09/866,028
;  PRIOR FILING DATE: 2001-05-25
;  NUMBER OF SEQ ID NOS: 120
; SEQ ID NO 42
;   LENGTH: 243
;   TYPE: PRT
```

```
;   ORGANISM: Homo Sapien
US-09-944-929-42

  Query Match             100.0%;  Score 1325;  DB 9;  Length 243;
  Best Local Similarity   100.0%;  Pred. No. 2.3e-89;
  Matches  243;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MRPLLVLLLLGLAAGSPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MRPLLVLLLLGLAAGSPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPG 60

Qy         61 APGEKGEGGRPGLPGPRGDPGPRGEAGPAGPTGPAGECSVPPRSAFSAKRSESRVPPPSD 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 APGEKGEGGRPGLPGPRGDPGPRGEAGPAGPTGPAGECSVPPRSAFSAKRSESRVPPPSD 120

Qy        121 APLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGESIASFFQ 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 APLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGESIASFFQ 180

Qy        181 FFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWHSSP 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 FFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWHSSP 240

Qy        241 VFA 243
              |||
Db        241 VFA 243


RESULT 7
US-10-028-072-362
; Sequence 362, Application US/10028072
; Publication No. US20030004311A1
; GENERAL INFORMATION:
;  APPLICANT: Baker,Kevin P.
;  APPLICANT:  Beresini,Maureen
;  APPLICANT:  DeForge,Laura
;  APPLICANT:  Desnoyers,Luc
;  APPLICANT:  Filvaroff,Ellen
;  APPLICANT:  Gao,Wei-Qiang
;  APPLICANT:  Gerritsen,Mary E.
;  APPLICANT:  Goddard,Audrey
;  APPLICANT:  Godowski,Paul J.
;  APPLICANT:  Gurney,Austin L.
;  APPLICANT:  Sherwood,Steven
;  APPLICANT:  Smith,Victoria
;  APPLICANT:  Stewart,Timothy A.
;  APPLICANT:  Tumas,Daniel
;  APPLICANT:  Watanabe,Colin K
;  APPLICANT:  Wood,William
;  APPLICANT:  Zhang
;  TITLE OF INVENTION:
;  FILE REFERENCE:
;  CURRENT APPLICATION NUMBER: US/10/028,072
;  CURRENT FILING DATE:  2001-12-19
;  PRIOR APPLICATION NUMBER: 60/049911
;  PRIOR FILING DATE: 1997-06-18
;  PRIOR APPLICATION NUMBER: 60/056974
;  PRIOR FILING DATE: 1997-08-26
;  PRIOR APPLICATION NUMBER: 60/059113
;  PRIOR FILING DATE: 1997-09-17
;  PRIOR APPLICATION NUMBER: 60/059115
;  PRIOR FILING DATE: 1997-09-17
;  PRIOR APPLICATION NUMBER: 60/059117
;  PRIOR FILING DATE: 1997-09-17
;  PRIOR APPLICATION NUMBER: 60/059122
;  PRIOR FILING DATE: 1997-09-17
;  PRIOR APPLICATION NUMBER: 60/059184
;  PRIOR FILING DATE: 1997-09-17
;  PRIOR APPLICATION NUMBER: 60/059263
;  PRIOR FILING DATE: 1997-09-18
;  PRIOR APPLICATION NUMBER: 60/059352
;  PRIOR FILING DATE: 1997-09-19
;  PRIOR APPLICATION NUMBER: 60/059588
;  PRIOR FILING DATE: 1997-09-19
;  PRIOR APPLICATION NUMBER: 60/059836
;  PRIOR FILING DATE: 1997-09-24
;  PRIOR APPLICATION NUMBER: 60/062250
;  PRIOR FILING DATE: 1997-10-17
;  PRIOR APPLICATION NUMBER: 60/062285
;  PRIOR FILING DATE: 1997-10-17
;  PRIOR APPLICATION NUMBER: 60/062287
;  PRIOR FILING DATE: 1997-10-17
;  PRIOR APPLICATION NUMBER: 60/062814
;  PRIOR FILING DATE: 1997-10-24
;  PRIOR APPLICATION NUMBER: 60/062816
;  PRIOR FILING DATE: 1997-10-24
;  PRIOR APPLICATION NUMBER: 60/063045
;  PRIOR FILING DATE: 1997-10-24
;  PRIOR APPLICATION NUMBER: 60/063082
;  PRIOR FILING DATE: 1997-10-31
;  PRIOR APPLICATION NUMBER: 60/063127
;  PRIOR FILING DATE: 1997-10-24
;  PRIOR APPLICATION NUMBER: 60/063327
;  PRIOR FILING DATE: 1997-10-27
;  PRIOR APPLICATION NUMBER: 60/063329
;  PRIOR FILING DATE: 1997-10-27
;  PRIOR APPLICATION NUMBER: 60/063550
;  PRIOR FILING DATE: 1997-10-28
;  PRIOR APPLICATION NUMBER: 60/063561
;  PRIOR FILING DATE: 1997-10-28
;  PRIOR APPLICATION NUMBER: 60/063704
;  PRIOR FILING DATE: 1997-10-29
;  PRIOR APPLICATION NUMBER: 60/063733
;  PRIOR FILING DATE: 1997-10-29
;  PRIOR APPLICATION NUMBER: 60/063735
;  PRIOR FILING DATE: 1997-10-29
;  PRIOR APPLICATION NUMBER: 60/063738
;  PRIOR FILING DATE: 1997-10-29
;  PRIOR APPLICATION NUMBER: 60/063755
;  PRIOR FILING DATE: 1997-10-17
;  PRIOR APPLICATION NUMBER: 60/064248
;  PRIOR FILING DATE: 1997-11-03
;  PRIOR APPLICATION NUMBER: 60/064809
;  PRIOR FILING DATE: 1997-11-07
;  PRIOR APPLICATION NUMBER: 60/065186
;  PRIOR FILING DATE: 1997-11-12
;  PRIOR APPLICATION NUMBER: 60/065846
;  PRIOR FILING DATE: 1997-11-17
;  PRIOR APPLICATION NUMBER: 60/066364
;  PRIOR FILING DATE: 1997-11-21
;  PRIOR APPLICATION NUMBER: 60/066453
;  PRIOR FILING DATE: 1997-11-24
;  PRIOR APPLICATION NUMBER: 60/066511
;  PRIOR FILING DATE: 1997-11-24
;  PRIOR APPLICATION NUMBER: 60/066770
;  PRIOR FILING DATE: 1997-11-24
;  PRIOR APPLICATION NUMBER: 60/069212
;  PRIOR FILING DATE: 1997-12-11
;  PRIOR APPLICATION NUMBER: 60/069278
;  PRIOR FILING DATE: 1997-12-11
;  PRIOR APPLICATION NUMBER: 60/069334
;  PRIOR FILING DATE: 1997-12-11
;  PRIOR APPLICATION NUMBER: 60/069694
;  PRIOR FILING DATE: 1997-12-16
;  PRIOR APPLICATION NUMBER: 60/072320
;  PRIOR FILING DATE: 1998-01-23
;  PRIOR APPLICATION NUMBER: 60/073612
;  PRIOR FILING DATE: 1998-02-04
;  PRIOR APPLICATION NUMBER: 60/074086
;  PRIOR FILING DATE: 1998-02-09
;  PRIOR APPLICATION NUMBER: 60/074092
;  PRIOR FILING DATE: 1998-02-09
;  PRIOR APPLICATION NUMBER: 60/077791
;  PRIOR FILING DATE: 1998-03-12
;  PRIOR APPLICATION NUMBER: 60/078910
;  PRIOR FILING DATE: 1998-03-20
```

; PRIOR APPLICATION NUMBER: 60/079294
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: 60/079663
; PRIOR FILING DATE: 1998-02-27
; PRIOR APPLICATION NUMBER: 60/079728
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/080165
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/081203
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: 60/081229
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: 60/081695
; PRIOR FILING DATE: 1998-04-14
; PRIOR APPLICATION NUMBER: 60/081817
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081818
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/082999
; PRIOR FILING DATE: 1998-04-24
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/083545
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084627
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084637
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/085149
; PRIOR FILING DATE: 1998-05-12
; PRIOR APPLICATION NUMBER: 60/085323
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085338
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085339
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085579
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085697
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085704
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/086414
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/086430
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/087106
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: 60/088026
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088730
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088741
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088810
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088858
; PRIOR FILING DATE: 19/98-06-11
; PRIOR APPLICATION NUMBER: 60/089532
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089599
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089907
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089947
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/090349
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090429
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090445
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090538
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090863
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/091360
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091519
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091982
; PRIOR FILING DATE: 1998-07-07

Query Match 100.0%; Score 1325; DB 9; Length 243;
Best Local Similarity 100.0%; Pred. No. 2.3e-89;
Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MRPLLVLLLLGLAAGSPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MRPLLVLLLLGLAAGSPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPG 60

Qy         61 APGEKGEGGRPGLPGPRGDPGPRGEAGPAGPTGPAGECSVPPRSAFSAKRSESRVPPPSD 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 APGEKGEGGRPGLPGPRGDPGPRGEAGPAGPTGPAGECSVPPRSAFSAKRSESRVPPPSD 120

Qy        121 APLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGESIASFFQ 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 APLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGESIASFFQ 180

Qy        181 FFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWHSSP 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 FFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWHSSP 240

Qy        241 VFA 243
              |||
Db        241 VFA 243
```

RESULT 8
US-10-121-049-362
; Sequence 362, Application US/10121049
; Publication No. US20030022239A1
; GENERAL INFORMATION:
; APPLICANT: Baker,Kevin P.
; APPLICANT: Beresini,Maureen
; APPLICANT: DeForge,Laura
; APPLICANT: Desnoyers,Luc
; APPLICANT: Filvaroff,Ellen
; APPLICANT: Gao,Wei-Qiang
; APPLICANT: Gerritsen,Mary E.
; APPLICANT: Goddard,Audrey
; APPLICANT: Godowski,Paul J.
; APPLICANT: Gurney,Austin L.
; APPLICANT: Sherwood,Steven
; APPLICANT: Smith,Victoria
; APPLICANT: Stewart,Timothy A.
; APPLICANT: Tumas,Daniel
; APPLICANT: Watanabe,Colin K
; APPLICANT: Wood,William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3330R1C17
; CURRENT APPLICATION NUMBER: US/10/121,049
; CURRENT FILING DATE: 2002-04-12
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 362
; LENGTH: 243
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-121-049-362

Query Match 100.0%; Score 1325; DB 9; Length 243;

```
  Best Local Similarity   100.0%;  Pred. No. 2.3e-89;
  Matches  243;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MRPLLVLLLLGLAAGSPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MRPLLVLLLLGLAAGSPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPG 60

Qy         61 APGEKGEGGRPGLPGPRGDPGPRGEAGPAGPTGPAGECSVPPRSAFSAKRSESRVPPPSD 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 APGEKGEGGRPGLPGPRGDPGPRGEAGPAGPTGPAGECSVPPRSAFSAKRSESRVPPPSD 120

Qy        121 APLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGESIASFFQ 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 APLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGESIASFFQ 180

Qy        181 FFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWHSSP 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 FFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWHSSP 240

Qy        241 VFA 243
              |||
Db        241 VFA 243


RESULT 9
US-10-123-904-362
; Sequence 362, Application US/10123904
; Publication No. US20030022328A1
; GENERAL INFORMATION:
;  APPLICANT: Baker,Kevin P.
;  APPLICANT:  Beresini,Maureen
;  APPLICANT:  DeForge,Laura
;  APPLICANT:  Desnoyers,Luc
;  APPLICANT:  Filvaroff,Ellen
;  APPLICANT:  Gao,Wei-Qiang
;  APPLICANT:  Gerritsen,Mary E.
;  APPLICANT:  Goddard,Audrey
;  APPLICANT:  Godowski,Paul J.
;  APPLICANT:  Gurney,Austin L.
;  APPLICANT:  Sherwood,Steven
;  APPLICANT:  Smith,Victoria
;  APPLICANT:  Stewart,Timothy A.
;  APPLICANT:  Tumas,Daniel
;  APPLICANT:  Watanabe,Colin K
;  APPLICANT:  Wood,William
;  APPLICANT:  Zhang, Zemin
;  TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;  TITLE OF INVENTION:  ACIDS ENCODING THE SAME
;  FILE REFERENCE: P3330R1C54
;  CURRENT APPLICATION NUMBER: US/10/123,904
;  CURRENT FILING DATE:  2002-04-16
;  Prior Application removed - See File Wrapper or Palm
;  NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 362
;   LENGTH: 243
;   TYPE: PRT
;   ORGANISM: Homo Sapien
US-10-123-904-362

  Query Match             100.0%;  Score 1325;  DB 9;  Length 243;
  Best Local Similarity   100.0%;  Pred. No. 2.3e-89;
  Matches  243;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MRPLLVLLLLGLAAGSPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MRPLLVLLLLGLAAGSPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPG 60

Qy         61 APGEKGEGGRPGLPGPRGDPGPRGEAGPAGPTGPAGECSVPPRSAFSAKRSESRVPPPSD 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 APGEKGEGGRPGLPGPRGDPGPRGEAGPAGPTGPAGECSVPPRSAFSAKRSESRVPPPSD 120

Qy        121 APLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGESIASFFQ 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 APLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGESIASFFQ 180

Qy        181 FFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWHSSP 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 FFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWHSSP 240

Qy        241 VFA 243
              |||
Db        241 VFA 243


RESULT 10
US-10-140-470-362
; Sequence 362, Application US/10140470
; Publication No. US20030022331A1
; GENERAL INFORMATION:
;  APPLICANT: Baker,Kevin P.
;  APPLICANT:  Beresini,Maureen
;  APPLICANT:  DeForge,Laura
;  APPLICANT:  Desnoyers,Luc
;  APPLICANT:  Filvaroff,Ellen
;  APPLICANT:  Gao,Wei-Qiang
;  APPLICANT:  Gerritsen,Mary E.
;  APPLICANT:  Goddard,Audrey
;  APPLICANT:  Godowski,Paul J.
;  APPLICANT:  Gurney,Austin L.
;  APPLICANT:  Sherwood,Steven
;  APPLICANT:  Smith,Victoria
;  APPLICANT:  Stewart,Timothy A.
;  APPLICANT:  Tumas,Daniel
;  APPLICANT:  Watanabe,Colin K
;  APPLICANT:  Wood,William
;  APPLICANT:  Zhang, Zemin
;  TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;  TITLE OF INVENTION:  ACIDS ENCODING THE SAME
;  FILE REFERENCE: P3330R1C160
;  CURRENT APPLICATION NUMBER: US/10/140,470
;  CURRENT FILING DATE:  2002-05-06
;  Prior Application removed - See Palm or File Wrapper
;  NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 362
;   LENGTH: 243
;   TYPE: PRT
;   ORGANISM: Homo Sapien
US-10-140-470-362

  Query Match             100.0%;  Score 1325;  DB 9;  Length 243;
  Best Local Similarity   100.0%;  Pred. No. 2.3e-89;
  Matches  243;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MRPLLVLLLLGLAAGSPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MRPLLVLLLLGLAAGSPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPG 60

Qy         61 APGEKGEGGRPGLPGPRGDPGPRGEAGPAGPTGPAGECSVPPRSAFSAKRSESRVPPPSD 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 APGEKGEGGRPGLPGPRGDPGPRGEAGPAGPTGPAGECSVPPRSAFSAKRSESRVPPPSD 120

Qy        121 APLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGESIASFFQ 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 APLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGESIASFFQ 180

Qy        181 FFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWHSSP 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 FFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWHSSP 240

Qy        241 VFA 243
              |||
Db        241 VFA 243


RESULT 11
US-09-796-753-68
```

; Sequence 68, Application US/09796753
; Publication No. US20030027998A1
; GENERAL INFORMATION:
; APPLICANT: McCarthy, Sean A.
; TITLE OF INVENTION: SECRETED PROTEINS AND USES THEREOF
; FILE REFERENCE: 7853-227-999
; CURRENT APPLICATION NUMBER: US/09/796,753
; CURRENT FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 09/183,175
; PRIOR FILING DATE: 1998-10-30
; PRIOR APPLICATION NUMBER: 09/223,094
; PRIOR FILING DATE: 1998-12-30
; PRIOR APPLICATION NUMBER: 09/223,546
; PRIOR FILING DATE: 1998-12-30
; PRIOR APPLICATION NUMBER: 09/224,246
; PRIOR FILING DATE: 1998-12-30
; PRIOR APPLICATION NUMBER: 09/259,388
; PRIOR FILING DATE: 1999-02-26
; PRIOR APPLICATION NUMBER: 60/122,458
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: 09/312,359
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: 09/336,536
; PRIOR FILING DATE: 1999-06-18
; PRIOR APPLICATION NUMBER: 09/342,687
; PRIOR FILING DATE: 1999-06-29
; PRIOR APPLICATION NUMBER: 09/345,464
; PRIOR FILING DATE: 1999-06-30
; PRIOR APPLICATION NUMBER: 09/365,164
; PRIOR FILING DATE: 1999-07-30
; PRIOR APPLICATION NUMBER: 09/399,723
; PRIOR FILING DATE: 1999-09-20
; PRIOR APPLICATION NUMBER: 09/409,634
; PRIOR FILING DATE: 1999-09-30
; PRIOR APPLICATION NUMBER: 09/471,179
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: 09/474,071
; PRIOR FILING DATE: 1999-12-29
; PRIOR APPLICATION NUMBER: 09/474,072
; PRIOR FILING DATE: 1999-12-29
; PRIOR APPLICATION NUMBER: 09/514,010
; PRIOR FILING DATE: 2000-02-25
; PRIOR APPLICATION NUMBER: 09/516,745
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 09/572,002
; PRIOR FILING DATE: 2000-05-14
; PRIOR APPLICATION NUMBER: 09/597,993
; PRIOR FILING DATE: 2000-06-19
; PRIOR APPLICATION NUMBER: 09/599,596
; PRIOR FILING DATE: 2000-06-22
; PRIOR APPLICATION NUMBER: 09/630,334
; PRIOR FILING DATE: 2000-07-31
; PRIOR APPLICATION NUMBER: 09/606,565
; PRIOR FILING DATE: 2000-06-29
; PRIOR APPLICATION NUMBER: 09/606,317
; PRIOR FILING DATE: 2000-06-29
; PRIOR APPLICATION NUMBER: 09/665,666
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: 09/677,751
; PRIOR FILING DATE: 2000-09-30
; NUMBER OF SEQ ID NOS: 162
; SEQ ID NO 68
; LENGTH: 243
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-796-753-68

Query Match 100.0%; Score 1325; DB 9; Length 243;
Best Local Similarity 100.0%; Pred. No. 2.3e-89;
Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MRPLLVLLLLGLAAGSPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPG 60
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1 MRPLLVLLLLGLAAGSPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPG 60

Qy 61 APGEKGEGGRPGLPGPRGDPGPRGEAGPAGPTGPAGECSVPPRSAFSAKRSESRVPPPSD 120
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 61 APGEKGEGGRPGLPGPRGDPGPRGEAGPAGPTGPAGECSVPPRSAFSAKRSESRVPPPSD 120

Qy 121 APLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGESIASFFQ 180
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 121 APLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGESIASFFQ 180

Qy 181 FFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWHSSP 240
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 181 FFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWHSSP 240

Qy 241 VFA 243
|||
Db 241 VFA 243

RESULT 12
US-10-175-746-362
; Sequence 362, Application US/10175746
; Publication No. US20030027270A1
; GENERAL INFORMATION:
; APPLICANT: Baker,Kevin P.
; APPLICANT: Beresini,Maureen
; APPLICANT: DeForge,Laura
; APPLICANT: Desnoyers,Luc
; APPLICANT: Filvaroff,Ellen
; APPLICANT: Gao,Wei-Qiang
; APPLICANT: Gerritsen,Mary E.
; APPLICANT: Goddard,Audrey
; APPLICANT: Godowski,Paul J.
; APPLICANT: Gurney,Austin L.
; APPLICANT: Sherwood,Steven
; APPLICANT: Smith,Victoria
; APPLICANT: Stewart,Timothy A.
; APPLICANT: Tumas,Daniel
; APPLICANT: Watanabe,Colin K
; APPLICANT: Wood,William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3330R1C353
; CURRENT APPLICATION NUMBER: US/10/175,746
; CURRENT FILING DATE: 2002-06-19
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 362
; LENGTH: 243
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-175-746-362

Query Match 100.0%; Score 1325; DB 9; Length 243;
Best Local Similarity 100.0%; Pred. No. 2.3e-89;
Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MRPLLVLLLLGLAAGSPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPG 60
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1 MRPLLVLLLLGLAAGSPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPG 60

Qy 61 APGEKGEGGRPGLPGPRGDPGPRGEAGPAGPTGPAGECSVPPRSAFSAKRSESRVPPPSD 120
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 61 APGEKGEGGRPGLPGPRGDPGPRGEAGPAGPTGPAGECSVPPRSAFSAKRSESRVPPPSD 120

Qy 121 APLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGESIASFFQ 180
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 121 APLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGESIASFFQ 180

Qy 181 FFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWHSSP 240
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 181 FFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWHSSP 240

```
Qy        241 VFA 243
              |||
Db        241 VFA 243


RESULT 13
US-10-176-918-362
; Sequence 362, Application US/10176918
; Publication No. US20030027275A1
; GENERAL INFORMATION:
;  APPLICANT: Baker,Kevin P.
;  APPLICANT:  Beresini,Maureen
;  APPLICANT:  DeForge,Laura
;  APPLICANT:  Desnoyers,Luc
;  APPLICANT:  Filvaroff,Ellen
;  APPLICANT:  Gao,Wei-Qiang
;  APPLICANT:  Gerritsen,Mary E.
;  APPLICANT:  Goddard,Audrey
;  APPLICANT:  Godowski,Paul J.
;  APPLICANT:  Gurney,Austin L.
;  APPLICANT:  Sherwood,Steven
;  APPLICANT:  Smith,Victoria
;  APPLICANT:  Stewart,Timothy A.
;  APPLICANT:  Tumas,Daniel
;  APPLICANT:  Watanabe,Colin K
;  APPLICANT:  Wood,William
;  APPLICANT:  Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION:  ACIDS ENCODING THE SAME
; FILE REFERENCE: P3330R1C382
; CURRENT APPLICATION NUMBER: US/10/176,918
; CURRENT FILING DATE:  2002-06-20
;     Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 362
;   LENGTH: 243
;   TYPE: PRT
;   ORGANISM: Homo Sapien
US-10-176-918-362

  Query Match             100.0%;  Score 1325;  DB 9;  Length 243;
  Best Local Similarity   100.0%;  Pred. No. 2.3e-89;
  Matches  243;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MRPLLVLLLLGLAAGSPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MRPLLVLLLLGLAAGSPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPG 60

Qy         61 APGEKGEGGRPGLPGPRGDPGPRGEAGPAGPTGPAGECSVPPRSAFSAKRSESRVPPPSD 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 APGEKGEGGRPGLPGPRGDPGPRGEAGPAGPTGPAGECSVPPRSAFSAKRSESRVPPPSD 120

Qy        121 APLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGESIASFFQ 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 APLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGESIASFFQ 180

Qy        181 FFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWHSSP 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 FFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWHSSP 240

Qy        241 VFA 243
              |||
Db        241 VFA 243


RESULT 14
US-10-176-921-362
; Sequence 362, Application US/10176921
; Publication No. US20030027276A1
; GENERAL INFORMATION:
;  APPLICANT: Baker,Kevin P.
;  APPLICANT:  Beresini,Maureen
;  APPLICANT:  DeForge,Laura
;  APPLICANT:  Desnoyers,Luc
;  APPLICANT:  Filvaroff,Ellen
;  APPLICANT:  Gao,Wei-Qiang
;  APPLICANT:  Gerritsen,Mary E.
;  APPLICANT:  Goddard,Audrey
;  APPLICANT:  Godowski,Paul J.
;  APPLICANT:  Gurney,Austin L.
;  APPLICANT:  Sherwood,Steven
;  APPLICANT:  Smith,Victoria
;  APPLICANT:  Stewart,Timothy A.
;  APPLICANT:  Tumas,Daniel
;  APPLICANT:  Watanabe,Colin K
;  APPLICANT:  Wood,William
;  APPLICANT:  Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION:  ACIDS ENCODING THE SAME
; FILE REFERENCE: P3330R1C288
; CURRENT APPLICATION NUMBER: US/10/176,921
; CURRENT FILING DATE:  2002-06-20
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 362
;   LENGTH: 243
;   TYPE: PRT
;   ORGANISM: Homo Sapien
US-10-176-921-362

  Query Match             100.0%;  Score 1325;  DB 9;  Length 243;
  Best Local Similarity   100.0%;  Pred. No. 2.3e-89;
  Matches  243;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MRPLLVLLLLGLAAGSPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MRPLLVLLLLGLAAGSPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPG 60

Qy         61 APGEKGEGGRPGLPGPRGDPGPRGEAGPAGPTGPAGECSVPPRSAFSAKRSESRVPPPSD 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 APGEKGEGGRPGLPGPRGDPGPRGEAGPAGPTGPAGECSVPPRSAFSAKRSESRVPPPSD 120

Qy        121 APLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGESIASFFQ 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 APLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGESIASFFQ 180

Qy        181 FFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWHSSP 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 FFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWHSSP 240

Qy        241 VFA 243
              |||
Db        241 VFA 243


RESULT 15
US-10-137-865-362
; Sequence 362, Application US/10137865
; Publication No. US20030032155A1
; GENERAL INFORMATION:
;  APPLICANT: Baker,Kevin P.
;  APPLICANT:  Beresini,Maureen
;  APPLICANT:  DeForge,Laura
;  APPLICANT:  Desnoyers,Luc
;  APPLICANT:  Filvaroff,Ellen
;  APPLICANT:  Gao,Wei-Qiang
;  APPLICANT:  Gerritsen,Mary E.
;  APPLICANT:  Goddard,Audrey
;  APPLICANT:  Godowski,Paul J.
;  APPLICANT:  Gurney,Austin L.
;  APPLICANT:  Sherwood,Steven
;  APPLICANT:  Smith,Victoria
;  APPLICANT:  Stewart,Timothy A.
;  APPLICANT:  Tumas,Daniel
;  APPLICANT:  Watanabe,Colin K
```

```
;  APPLICANT:  Wood,William
;  APPLICANT:  Zhang, Zemin
;  TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;  TITLE OF INVENTION:  ACIDS ENCODING THE SAME
;  FILE REFERENCE: P3330R1C154
;  CURRENT APPLICATION NUMBER: US/10/137,865
;  CURRENT FILING DATE:  2002-05-03
;  Prior Application removed - See Palm or File Wrapper
;  NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 362
;   LENGTH: 243
;   TYPE: PRT
;   ORGANISM: Homo Sapien
US-10-137-865-362

  Query Match           100.0%;  Score 1325;  DB 9;  Length 243;
  Best Local Similarity  100.0%;  Pred. No. 2.3e-89;
  Matches  243;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MRPLLVLLLLGLAAGSPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MRPLLVLLLLGLAAGSPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPG 60

Qy         61 APGEKGEGGRPGLPGPRGDPGPRGEAGPAGPTGPAGECSVPPRSAFSAKRSESRVPPPSD 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 APGEKGEGGRPGLPGPRGDPGPRGEAGPAGPTGPAGECSVPPRSAFSAKRSESRVPPPSD 120

Qy        121 APLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGESIASFFQ 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 APLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGESIASFFQ 180

Qy        181 FFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWHSSP 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 FFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWHSSP 240

Qy        241 VFA 243
              |||
Db        241 VFA 243


Search completed: June 18, 2003, 15:02:35
Job time : 57 secs
```

GenCore version 5.1.6


OM protein - protein search, using sw model

Run on: June 18, 2003, 14:54:03 ; Search time 71 Seconds
(without alignments)
456.055 Million cell updates/sec

Title: US-09-943-851A-42
Perfect score: 243
Sequence: 1 MRPLLVLLLLGLAAGSPPLD.....DSTFSGFLVYSDWHSSPVFA 243

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 908470 seqs, 133250620 residues

Word size : 0

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : A_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 243 | 100.0 | 243 | 20 | AAY06481 | Human tumour-assoc |
| 2 | 243 | 100.0 | 243 | 20 | AAY17827 | Human PRO344 prote |
| 3 | 243 | 100.0 | 243 | 20 | AAW97984 | Human adipocyte-sp |
| 4 | 243 | 100.0 | 243 | 21 | AAB33461 | Human PRO344 prote |
| 5 | 243 | 100.0 | 243 | 21 | AAY71468 | Human PRO344 prote |
| 6 | 243 | 100.0 | 243 | 21 | AAY93688 | Amino acid sequenc |
| 7 | 243 | 100.0 | 243 | 21 | AAB01318 | Human PRO344 polyp |
| 8 | 243 | 100.0 | 243 | 22 | AAU12352 | Human PRO344 polyp |
| 9 | 243 | 100.0 | 243 | 22 | AAB65815 | Human TANGO 253 SE |
| 10 | 243 | 100.0 | 243 | 22 | AAB49593 | Human adipocyte co |
| 11 | 243 | 100.0 | 243 | 22 | AAB49599 | Human adipocyte co |
| 12 | 228 | 93.8 | 228 | 22 | AAB65816 | Human mature TANGO |
| 13 | 213 | 87.7 | 220 | 22 | ABG12724 | Novel human diagno |
| 14 | 185 | 76.1 | 243 | 22 | AAB65888 | Human secreted pro |
| 15 | 174 | 71.6 | 243 | 22 | AAB65891 | Human secreted pro |
| 16 | 173 | 71.2 | 201 | 23 | AAO21663 | Human secreted pro |
| 17 | 148 | 60.9 | 243 | 22 | AAB65889 | Human secreted pro |
| 18 | 142 | 58.4 | 243 | 22 | AAB65890 | Human secreted pro |
| 19 | 128 | 52.7 | 128 | 22 | AAB65819 | Human TANGO 253 C1 |
| 20 | 82 | 33.7 | 151 | 22 | ABG12723 | Novel human diagno |
| 21 | 70 | 28.8 | 243 | 21 | AAY76040 | Rat skin cell prot |
| 22 | 70 | 28.8 | 243 | 22 | AAB55979 | Skin cell protein, |
| 23 | 70 | 28.8 | 243 | 23 | ABB72179 | Rat protein isolat |
| 24 | 62 | 25.5 | 228 | 22 | AAB65821 | Murine mature TANG |
| 25 | 62 | 25.5 | 242 | 22 | AAB65896 | Murine secreted pr |
| 26 | 62 | 25.5 | 243 | 22 | AAB65820 | Murine TANGO 253 S |
| 27 | 62 | 25.5 | 243 | 22 | AAB65897 | Murine secreted pr |
| 28 | 62 | 25.5 | 243 | 22 | AAB65898 | Murine secreted pr |
| 29 | 62 | 25.5 | 243 | 22 | AAB65899 | Murine secreted pr |
| 30 | 60 | 24.7 | 60 | 22 | AAB65818 | Human TANGO 253 co |
| 31 | 51 | 21.0 | 128 | 22 | AAB65824 | Murine TANGO 253 C |
| 32 | 43 | 17.7 | 60 | 22 | AAB65823 | Murine TANGO 253 c |
| 33 | 30 | 12.3 | 87 | 21 | AAY87269 | Human signal pepti |
| 34 | 15 | 6.2 | 15 | 22 | AAB65817 | Human TANGO 253 si |
| 35 | 11 | 4.5 | 27 | 22 | AAB92425 | Miscellaneous pept |
| 36 | 11 | 4.5 | 56 | 22 | ABB31339 | Peptide #3990 enco |
| 37 | 11 | 4.5 | 56 | 22 | ABB36545 | Peptide #4051 enco |
| 38 | 11 | 4.5 | 56 | 22 | ABB21884 | Protein #3883 enco |
| 39 | 11 | 4.5 | 56 | 22 | AAM57307 | Human brain expres |
| 40 | 11 | 4.5 | 56 | 22 | AAM69711 | Human bone marrow |
| 41 | 11 | 4.5 | 56 | 22 | AAM17524 | Peptide #3958 enco |
| 42 | 11 | 4.5 | 56 | 22 | AAM30045 | Peptide #4082 enco |
| 43 | 11 | 4.5 | 56 | 22 | AAM05196 | Peptide #3878 enco |
| 44 | 11 | 4.5 | 56 | 23 | ABG39329 | Human peptide enco |
| 45 | 11 | 4.5 | 128 | 20 | AAY02690 | Human secreted pro |

ALIGNMENTS

RESULT 1
AAY06481
ID AAY06481 standard; Protein; 243 AA.
XX
AC AAY06481;
XX
DT 27-SEP-1999 (first entry)
XX
DE Human tumour-associated protein PRO344.
XX
KW PRO344; UNQ303; cancer; tumour; diagnosis; therapy; human.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Peptide 1..15
FT /note= "signal peptide"
FT Protein 16..243
FT /note= "mature protein"
FT Modified-site 68..215
FT /note= "N-myristoylated"
FT Modified-site 216..243
FT /note= "N-myristoylated"
XX
PN WO9935170-A2.
XX
PD 15-JUL-1999.
XX
PF 05-JAN-1999; 99WO-US00106.
XX
PR 20-NOV-1998; 98US-0109304.
PR 05-JAN-1998; 98US-0070440.
PR 29-APR-1998; 98US-0083500.

PR 22-MAY-1998; 98US-0086414.
PR 10-JUN-1998; 98US-0088742.
PR 10-NOV-1998; 98US-0107783.
XX
PA (GETH ) GENENTECH INC.
XX
PI Botstein D, Goddard A, Gurney AL, Hillan KJ, Lawrence DA;
PI Roy MA, Wood WI;
XX
DR WPI; 1999-430385/36.
DR N-PSDB; AAX87258.
XX
PT Antibody against proteins expressed in neoplastic cells, useful for
PT tumor diagnosis and treatment
XX
PS Example 1; Fig 10; 162pp; English.
XX
CC This sequence represents human PRO344 (UNQ303), a protein encoded
CC by the novel cDNA clone DNA40592 (see AAX87258). Amplification of
CC DNA40592 was observed in primary lung tumours and in primary colon
CC tumours, suggesting a significant role in tumour formation and
CC growth. Antagonists (e.g. antibodies) directed to PRO344 may have
CC use in cancer therapy. The invention identifies 14 genes (see
CC AAX87254-67) that are amplified in the genome of tumour cells. Such
CC amplification is expected to be associated with overexpression of
CC the gene product and to contribute to tumorigenesis. The encoded
CC proteins (see AAY06477-90) may be useful targets for the diagnosis
CC and/or treatment (including prevention) of certain cancers, and may
CC act as predictors of the prognosis of tumour treatment. Antibodies
CC that bind the proteins are claimed and used in claimed cancer
CC diagnostic kits.
XX
SQ Sequence 243 AA;

Query Match 100.0%; Score 243; DB 20; Length 243;
Best Local Similarity 100.0%; Pred. No. 4.9e-212;
Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MRPLLVLLLLGLAAGSPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MRPLLVLLLLGLAAGSPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPG 60

Qy         61 APGEKGEGGRPGLPGPRGDPGPRGEAGPAGPTGPAGECSVPPRSAFSAKRSESRVPPPSD 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 APGEKGEGGRPGLPGPRGDPGPRGEAGPAGPTGPAGECSVPPRSAFSAKRSESRVPPPSD 120

Qy        121 APLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGESIASFFQ 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 APLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGESIASFFQ 180

Qy        181 FFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWHSSP 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 FFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWHSSP 240

Qy        241 VFA 243
              |||
Db        241 VFA 243
```

RESULT 2
AAY17827
ID AAY17827 standard; Protein; 243 AA.
XX
AC AAY17827;
XX
DT 12-AUG-1999 (first entry)
XX
DE Human PRO344 protein sequence.
XX
KW Human; PRO protein; tumour necrosis factor family; TNF; cytokine;
KW secreted protein; transmembrane protein; inflammation disorder.
XX
OS Homo sapiens.

XX
PN WO9928462-A2.
XX
PD 10-JUN-1999.
XX
PF 01-DEC-1998; 98WO-US25108.
XX
PR 25-FEB-1998; 98US-0075945.
PR 03-DEC-1997; 97US-0067411.
PR 11-DEC-1997; 97US-0069278.
PR 11-DEC-1997; 97US-0069334.
PR 11-DEC-1997; 97US-0069335.
PR 12-DEC-1997; 97US-0069425.
PR 16-DEC-1997; 97US-0069694.
PR 16-DEC-1997; 97US-0069696.
PR 16-DEC-1997; 97US-0069702.
PR 17-DEC-1997; 97US-0069870.
PR 17-DEC-1997; 97US-0069873.
PR 18-DEC-1997; 97US-0068017.
PR 05-JAN-1998; 98US-0070440.
PR 09-FEB-1998; 98US-0074086.
PR 09-FEB-1998; 98US-0074092.
XX
PA (GETH ) GENENTECH INC.
XX
PI Baker KP, Chen J, Goddard A, Gurney AL, Wood WI;
PI Yuan J;
XX
DR WPI; 1999-371118/31.
DR N-PSDB; AAX80052.
XX
PT Nucleic acids encoding PRO secreted and transmembrane proteins
XX
PS Claim 12; Fig 21; 123pp; English.
XX
CC The present invention describes nucleic acids encoding PRO secreted and
CC transmembrane proteins used therapeutically. The PRO proteins have
CC cytostatic, anti-inflammatory, anti-proliferative and immunosuppressive
CC activity. The proteins and polynucleotides can be used in therapy,
CC identification of homologues, raising antibodies and design of probes
CC and primers. They can be used in a range of diseases related to proteins
CC that they have homology with, e.g. a PRO protein having homology to
CC complement proteins may be used in inflammatory responses.
XX
SQ Sequence 243 AA;

Query Match 100.0%; Score 243; DB 20; Length 243;
Best Local Similarity 100.0%; Pred. No. 4.9e-212;
Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MRPLLVLLLLGLAAGSPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MRPLLVLLLLGLAAGSPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPG 60

Qy         61 APGEKGEGGRPGLPGPRGDPGPRGEAGPAGPTGPAGECSVPPRSAFSAKRSESRVPPPSD 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 APGEKGEGGRPGLPGPRGDPGPRGEAGPAGPTGPAGECSVPPRSAFSAKRSESRVPPPSD 120

Qy        121 APLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGESIASFFQ 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 APLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGESIASFFQ 180

Qy        181 FFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWHSSP 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 FFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWHSSP 240

Qy        241 VFA 243
              |||
Db        241 VFA 243
```

RESULT 3
AAW97984

```
ID   AAW97984 standard; Protein; 243 AA.
XX
AC   AAW97984;
XX
DT   21-JUN-1999  (first entry)
XX
DE   Human adipocyte-specific protein zsig39.
XX
KW   Adipocyte-specific protein; zsig39; human; fatty acid metabolism;
KW   energy balance; nutrition; antimicrobial.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1..15
FT                   /note= "signal peptide, alternatively the signal
FT                          peptide comprises residues 1..18"
FT   Protein         16..243
FT                   /note= "mature protein, alternatively the mature
FT                          protein comprises residues 19..243
FT                          (specifically claimed in Claim 4)"
FT   Domain          30..96
FT                   /note= "collagen-like domain"
FT   Domain          98..243
FT                   /note= "globular domain"
FT   Region          105..109
FT                   /note= "beta strand"
FT   Region          128..130
FT                   /note= "beta strand"
FT   Region          136..139
FT                   /note= "beta strand"
FT   Region          143..146
FT                   /note= "beta strand"
FT   Region          164..171
FT                   /note= "beta strand"
FT   Region          176..182
FT                   /note= "beta strand"
FT   Region          187..200
FT                   /note= "beta strand"
FT   Region          204..210
FT                   /note= "beta strand"
FT   Region          226..231
FT                   /note= "beta strand"
FT   Domain          111..135
FT                   /note= "receptor binding domain"
FT   Domain          170..174
FT                   /note= "receptor binding domain"
XX
PN   WO9910492-A1.
XX
PD   04-MAR-1999.
XX
PF   26-AUG-1998;   98WO-US17724.
XX
PR   26-AUG-1997;   97US-0056983.
XX
PA   (ZYMO ) ZYMOGENETICS INC.
XX
PI   Humes JM,  Sheppard PO;
XX
DR   WPI; 1999-204665/17.
DR   N-PSDB; AAX24684.
XX
PT   Zsig39 protein - used to modulate fatty acid metabolism
XX
PS   Claim 1; 111-112; 132pp; English.
XX
CC   This polypeptide comprises human adipocyte-specific protein zsig39,
CC   a protein that modulates free fatty acid metabolism.  zsig39 is a
CC   member of a family of proteins having a globular domain and a
CC   collagen-like domain capable of dimerisation or oligomerisation.
CC   zsig39 polypeptides were initially identified by querying an EST
CC   database for secretory signal sequences characterised by an upstream
CC   methionine start site, a hydrophobic region of approximately 13
CC   amino acids and a cleavage site.  A single EST sequence was
CC   discovered, and the novel polypeptide encoded by the full-length
CC   cDNA allowed the identification of a homologue relationship with
CC   adipocyte complement related protein Acrp30 and adipocyte secreted
CC   protein apM1.  A full-length clone (see AAX24684) was obtained from a
CC   lung tissue library.  Expression vectors, cultured cells and a
CC   method of producing zsig39 polypeptide are claimed, as well as
CC   zsig39 polypeptides having N- or C-terminal affinity tags, toxins,
CC   radionucleotides, enzymes or fluorophores, fusion proteins
CC   including zsig39 polypeptides, an antibody that specifically
CC   binds to an epitope of zsig39, and a method for modulating free
CC   fatty acid metabolism by administering a zsig39 polypeptide.  The
CC   zsig39 polypeptide may also be used in organ preservation, for
CC   cryopreservation, for surgical pretreatment to prevent injury due
CC   to ischaemia and/or inflammation, and as an antimicrobial agent,
CC   promoting lysis or phagocytosis of infectious agents.
XX
SQ   Sequence   243 AA;

  Query Match             100.0%;  Score 243;  DB 20;  Length 243;
  Best Local Similarity   100.0%;  Pred. No. 4.9e-212;
  Matches  243;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MRPLLVLLLLGLAAGSPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MRPLLVLLLLGLAAGSPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPG 60

Qy         61 APGEKGEGGRPGLPGPRGDPGPRGEAGPAGPTGPAGECSVPPRSAFSAKRSESRVPPPSD 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 APGEKGEGGRPGLPGPRGDPGPRGEAGPAGPTGPAGECSVPPRSAFSAKRSESRVPPPSD 120

Qy        121 APLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGESIASFFQ 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 APLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGESIASFFQ 180

Qy        181 FFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWHSSP 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 FFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWHSSP 240

Qy        241 VFA 243
              |||
Db        241 VFA 243


RESULT 4
AAB33461
ID   AAB33461 standard; Protein; 243 AA.
XX
AC   AAB33461;
XX
DT   29-JAN-2001  (first entry)
XX
DE   Human PRO344 protein UNQ303 SEQ ID NO:241.
XX
KW   Human; immune related disease; diagnosis; antiinflammatory; cardiant;
KW   dermatological; antiarthritic; antirheumatic; immunosuppressive;
KW   haemostatic; antithyroid; antidiabetic; nootropic; neuroprotective;
KW   antianaemic; hepatotropic; virucide; antipsoriatic; antiallergic;
KW   antiasthmatic; systemic lupus erythematosus; rheumatoid arthritis;
KW   osteoarthritis; spondyloarthropathy; systemic sclerosis; sarcoidosis;
KW   idiopathic inflammatory myopathy; Sjogren's syndrome; thyroiditis;
KW   systemic vasculitis; autoimmune haemolytic anaemia; diabetes mellitus;
KW   autoimmune thrombocytopaenia; immune-mediated renal disease;
KW   demyelinating disease; hepatobiliary disease; Whipple's disease;
KW   inflammatory bowel disease; gluten-sensitive enteropathy;
KW   autoimmune disease; immune-mediated skin disease; allergic disease;
KW   immunological disease; transplantation associated disease;
KW   graft rejection; graft-versus-host-disease.
XX
OS   Homo sapiens.
XX
PN   WO200053758-A2.
```

```
XX
PD   14-SEP-2000.
XX
PF   02-MAR-2000; 2000WO-US05841.
XX
PR   08-MAR-1999;   99WO-US05028.
PR   10-MAR-1999;   99US-0123618.
PR   12-MAR-1999;   99US-0123957.
PR   23-MAR-1999;   99US-0125775.
PR   12-APR-1999;   99US-0128849.
PR   20-APR-1999;   99WO-US08615.
PR   28-APR-1999;   99US-0131445.
PR   04-MAY-1999;   99US-0132371.
PR   14-MAY-1999;   99US-0134287.
PR   02-JUN-1999;   99WO-US12252.
PR   23-JUN-1999;   99US-0141037.
PR   20-JUL-1999;   99US-0144758.
PR   26-JUL-1999;   99US-0145698.
PR   28-JUL-1999;   99US-0146222.
PR   01-SEP-1999;   99WO-US20111.
PR   08-SEP-1999;   99WO-US20594.
PR   13-SEP-1999;   99WO-US20944.
PR   15-SEP-1999;   99WO-US21090.
PR   15-SEP-1999;   99WO-US21547.
PR   05-OCT-1999;   99WO-US23089.
PR   29-OCT-1999;   99US-0162506.
PR   29-NOV-1999;   99WO-US28214.
PR   30-NOV-1999;   99WO-US28313.
PR   30-NOV-1999;   99WO-US28409.
PR   01-DEC-1999;   99WO-US28301.
PR   01-DEC-1999;   99WO-US28634.
PR   02-DEC-1999;   99WO-US28551.
PR   02-DEC-1999;   99WO-US28564.
PR   02-DEC-1999;   99WO-US28565.
PR   16-DEC-1999;   99WO-US30095.
PR   20-DEC-1999;   99WO-US30099.
PR   30-DEC-1999;   99WO-US31274.
PR   05-JAN-2000; 2000WO-US00219.
PR   06-JAN-2000; 2000WO-US00277.
PR   06-JAN-2000; 2000WO-US00376.
PR   11-FEB-2000; 2000WO-US03565.
PR   18-FEB-2000; 2000WO-US04341.
PR   18-FEB-2000; 2000WO-US04342.
PR   22-FEB-2000; 2000WO-US04414.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Ashkenazi AJ,  Baker KP,  Goddard A,  Gurney AL,  Hebert C,  Henzel W;
PI   Kabakoff RC,  Lu Y,  Pan J,  Pennica D,  Shelton DL,  Smith V;
PI   Stewart TA,  Tumas D,  Watanabe CK,  Wood WI,  Yan M;
XX
DR   WPI; 2000-572271/53.
DR   N-PSDB; AAC58626.
XX
PT   Sixty four PRO polypeptides, useful in the diagnosis and treatment of
PT   immune related disorders, e.g. systemic lupus erythematosis, rheumatoid
PT   arthritis, osteoarthritis, thyroiditis and diabetes mellitus -
XX
PS   Claim 33; Fig 96; 309pp; English.
XX
CC   The present invention describes sixty four human PRO proteins which can
CC   be used in the treatment of immune related diseases. The human PRO
CC   proteins, anti-PRO antibodies, agonists and antagonists are useful for
CC   treating and diagnosing immune related disorders. The disorders are
CC   selected from systemic lupus erythematosus, rheumatoid arthritis,
CC   osteoarthritis, juvenile chronic arthritis, spondyloarthropathies,
CC   systemic sclerosis, idiopathic inflammatory myopathies, Sjogren's
CC   syndrome, systemic vasculitis, sarcoidosis, autoimmune haemolytic
CC   anaemia, autoimmune thrombocytopaenia, thyroiditis, diabetes mellitus,
CC   immune-mediated renal disease, demyelinating diseases of the central
CC   and peripheral nervous systems, hepatobiliary diseases, inflammatory
CC   bowel disease, gluten-sensitive enteropathy and Whipple's disease,
CC   autoimmune or immune-mediated skin diseases, allergic diseases,
CC   immunological diseases of the lung, and transplantation associated
CC   diseases including graft rejection and graft-versus-host-disease.
CC   AAC58397 to AAC58578 represent PCR primers and hybridisation probes used
CC   in the isolation of human PRO sequences. AAC58579 to AAC58642 and
CC   AAB33414 to AAB33477 represent human PRO polynucleotide and protein
CC   sequences given in the exemplification of the present invention.
XX
SQ   Sequence   243 AA;

  Query Match             100.0%;  Score 243;  DB 21;  Length 243;
  Best Local Similarity   100.0%;  Pred. No. 4.9e-212;
  Matches  243;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MRPLLVLLLLGLAAGSPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MRPLLVLLLLGLAAGSPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPG 60

Qy         61 APGEKGEGGRPGLPGPRGDPGPRGEAGPAGPTGPAGECSVPPRSAFSAKRSESRVPPPSD 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 APGEKGEGGRPGLPGPRGDPGPRGEAGPAGPTGPAGECSVPPRSAFSAKRSESRVPPPSD 120

Qy        121 APLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGESIASFFQ 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 APLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGESIASFFQ 180

Qy        181 FFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWHSSP 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 FFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWHSSP 240

Qy        241 VFA 243
              |||
Db        241 VFA 243


RESULT 5
AAY71468
ID   AAY71468 standard; Protein; 243 AA.
XX
AC   AAY71468;
XX
DT   08-NOV-2000  (first entry)
XX
DE   Human PRO344 protein.
XX
KW   PRO344; DNA40592-1242; human; ATCC No: 209492; antiproliferative;
KW   neoplastic cell growth inhibitor; cytostatic; treatment; cancer; tumour;
KW   breast; prostate; colon; lung; renal; ovarian; central nervous system;
KW   CNS; leukemia; melanoma; Expressed Sequence Tag; EST; secreted protein;
KW   extracellular domain; ECD.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1..15
FT                   /label= Signal_peptide
FT   Modified-site   11..17
FT                   /note= "N-myristoylation site"
FT   Protein         16..243
FT                   /label= Mature_PRO344_protein
FT   Modified-site   68..74
FT                   /note= "N-myristoylation site"
FT   Binding-site    77..80
FT                   /note= "Cell attachment sequence"
FT   Modified-site   216..222
FT                   /note= "N-myristoylation site"
XX
PN   WO200032778-A2.
XX
PD   08-JUN-2000.
XX
PF   30-NOV-1999;   99WO-US28409.
XX
PR   01-DEC-1998;   98WO-US25108.
```

```
PR   16-DEC-1998;   98US-0112850.
PR   22-DEC-1998;   98US-0113296.
PR   20-JUL-1999;   99US-0144758.
PR   26-JUL-1999;   99US-0145698.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Chen J,  Goddard A,  Gurney AL,  Hillan K,  Napier M,  Wood WI;
XX
DR   WPI; 2000-412325/35.
DR   N-PSDB; AAD01241.
XX
PT   New composition useful for inhibiting neoplastic cell growth and for
PT   treating cancers, comprises PRO655, PRO344 or PRO364 polypeptide or
PT   their antagonists -
XX
PS   Claim 31; Fig 6; 108pp; English.
XX
CC   The present sequence is the human PRO344 protein, encoded by the cDNA
CC   clone, designated as DNA40592-1242. It is isolated from human foetal
CC   lung tissue, cDNA library, identified using probes based on a consensus
CC   sequence DNA34398, derived from secreted protein extracellular domain
CC   (ECD) expressed sequence tag (EST). This clone is assigned ATCC deposit
CC   No: 209492. PRO344 functions as a neoplastic cell growth inhibitor and
CC   is used for treating tumours, using an effective amount of PRO655, PRO364
CC   and PRO344. This composition is especially useful for treatment of human
CC   cancers such as breast, prostate, colon, lung, renal, ovarian and CNS,
CC   leukemia and melanoma.
XX
SQ   Sequence   243 AA;

  Query Match             100.0%;  Score 243;  DB 21;  Length 243;
  Best Local Similarity   100.0%;  Pred. No. 4.9e-212;
  Matches  243;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MRPLLVLLLLGLAAGSPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MRPLLVLLLLGLAAGSPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPG 60

Qy         61 APGEKGEGGRPGLPGPRGDPGPRGEAGPAGPTGPAGECSVPPRSAFSAKRSESRVPPPSD 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 APGEKGEGGRPGLPGPRGDPGPRGEAGPAGPTGPAGECSVPPRSAFSAKRSESRVPPPSD 120

Qy        121 APLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGESIASFFQ 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 APLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGESIASFFQ 180

Qy        181 FFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWHSSP 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 FFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWHSSP 240

Qy        241 VFA 243
              |||
Db        241 VFA 243


RESULT 6
AAY93688
ID   AAY93688 standard; Protein; 243 AA.
XX
AC   AAY93688;
XX
DT   03-OCT-2000  (first entry)
XX
DE   Amino acid sequence of novel polypeptide PRO344.
XX
KW   PRO201; PRO292; PRO327; PRO1265;  PRO344; PRO343; PRO347; PRO357;
KW   PRO715; PRO1017; PRO1112; PRO509; PRO853; PRO882; tumour cell;
KW   tumourigenesis; cancer; neoplastic cell growth; cell proliferation.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers

FT   Peptide         1..15
FT                   /note= "signal peptide"
FT   Modified-site   11..17
FT                   /note= "N-myristoylation site"
FT   Modified-site   68..74
FT                   /note= "N-myristoylation site"
FT   Region          77..80
FT                   /note= "cell attachment sequence"
FT   Modified-site   216..222
FT                   /note= "N-myristoylation site"
XX
PN   WO200037640-A2.
XX
PD   29-JUN-2000.
XX
PF   16-DEC-1999;   99WO-US30095.
XX
PR   22-DEC-1998;   98US-0113296.
PR   08-MAR-1999;   99WO-US05028.
PR   02-JUN-1999;   99WO-US12252.
PR   01-SEP-1999;   99WO-US20111.
PR   15-SEP-1999;   99WO-US21090.
PR   30-NOV-1999;   99WO-US28313.
PR   30-NOV-1999;   99WO-US28409.
PR   01-DEC-1999;   99WO-US28301.
PR   02-DEC-1999;   99WO-US28565.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Botstein D,  Goddard A,  Gurney AL,  Hillan K,  Lawrence DA,  Roy MA;
PI   Wood WI;
XX
DR   WPI; 2000-452188/39.
DR   N-PSDB; AAA46907.
XX
PT   New anti-polypeptide antibody useful in the treatment and diagnosis of
PT   neoplastic cell growth and proliferation -
XX
PS   Claim 61; Fig 10; 220pp; English.
XX
CC   The present sequence represents a novel human polypeptide. The
CC   specification describes novel polypeptides designated PRO201, PRO292,
CC   PRO327, PRO1265, PRO344, PRO343, PRO347, PRO357, PRO715, PRO1017,
CC   PRO1112, PRO509, PRO853 and PRO882. These genes are amplified in
CC   the genome of tumour cells. The polypeptides are believed to contribute
CC   to tumourigenesis. The polypeptides are useful target for the
CC   identification of certain cancers, and may act as predictors of the
CC   prognosis of tumour treatment. Antibodies against these polypeptides
CC   are useful in the treatment and diagnosis of neoplastic cell growth
CC   and proliferation in mammals.
XX
SQ   Sequence   243 AA;

  Query Match             100.0%;  Score 243;  DB 21;  Length 243;
  Best Local Similarity   100.0%;  Pred. No. 4.9e-212;
  Matches  243;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MRPLLVLLLLGLAAGSPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MRPLLVLLLLGLAAGSPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPG 60

Qy         61 APGEKGEGGRPGLPGPRGDPGPRGEAGPAGPTGPAGECSVPPRSAFSAKRSESRVPPPSD 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 APGEKGEGGRPGLPGPRGDPGPRGEAGPAGPTGPAGECSVPPRSAFSAKRSESRVPPPSD 120

Qy        121 APLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGESIASFFQ 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 APLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGESIASFFQ 180

Qy        181 FFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWHSSP 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 FFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWHSSP 240
```

```
Qy        241 VFA 243
              |||
Db        241 VFA 243


RESULT 7
AAB01318
ID   AAB01318 standard; Protein; 243 AA.
XX
AC   AAB01318;
XX
DT   25-SEP-2000  (first entry)
XX
DE   Human PRO344 polypeptide.
XX
KW   PRO; membrane bound protein; secreted protein; PRO357; PRO327;
KW   PRO243; PRO715; PRO241; PRO323; PRO299; PRO233; PRO344; PRO347;
KW   PRO355; PRO353; PRO361; PRO365; transmembrane polypeptide;
KW   antibody; screening; detection; inhibition; probe; primer; human.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1..15
FT                   /label= Signal peptide
FT   Modified-site   11..17
FT                   /note= "N-myristoylation site"
FT   Modified-site   68..74
FT                   /note= "N-myristoylation site"
FT   Region          77..80
FT                   /label= Cell attachment sequence
FT   Modified-site   216..222
FT                   /note= "N-myristoylation site"
XX
PN   WO200032776-A2.
XX
PD   08-JUN-2000.
XX
PF   01-DEC-1999;   99WO-US28301.
XX
PR   01-DEC-1998;   98WO-US25108.
PR   16-DEC-1998;   98US-0112850.
PR   22-DEC-1998;   98US-0113296.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Baker KP,  Botstein D,  Eaton DL,  Ferrara N,  Filvaroff E;
PI   Gerritsen ME,  Goddard A,  Godowski PJ,  Grimaldi CJ,  Gurney AL;
PI   Hillan KJ,  Kljavin IJ,  Napier MA,  Roy MA,  Tumas D,  Wood WI;
XX
DR   WPI; 2000-412324/35.
DR   N-PSDB; AAA49560.
XX
PT   New human nucleic acids encoding secreted and transmembrane
PT   polypeptides, designated as PRO polypeptides, useful as pharmaceutical
PT   and diagnostic agents
XX
PS   Claim 12; Fig 18; 187pp; English.
XX
CC   New human nucleic acids encoding secreted and transmembrane
CC   polypeptides which are designated as PRO polypeptides are described
CC   The membrane-bound proteins have various industrial applications,
CC   including as pharmaceutical and diagnostic agents. The membrane-bound
CC   proteins can also be employed for screening of potential peptide or
CC   small molecule inhibitors of the relevant receptor/ligand interaction.
CC   Anti-PRO antibodies are useful for the affinity purification of PRO
CC   from recombinant cell culture or natural sources.
XX
SQ   Sequence   243 AA;

 Query Match           100.0%;  Score 243;  DB 21;  Length 243;
 Best Local Similarity  100.0%;  Pred. No. 4.9e-212;
 Matches  243;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MRPLLVLLLLGLAAGSPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MRPLLVLLLLGLAAGSPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPG 60

Qy         61 APGEKGEGGRPGLPGPRGDPGPRGEAGPAGPTGPAGECSVPPRSAFSAKRSESRVPPPSD 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 APGEKGEGGRPGLPGPRGDPGPRGEAGPAGPTGPAGECSVPPRSAFSAKRSESRVPPPSD 120

Qy        121 APLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGESIASFFQ 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 APLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGESIASFFQ 180

Qy        181 FFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWHSSP 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 FFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWHSSP 240

Qy        241 VFA 243
              |||
Db        241 VFA 243


RESULT 8
AAU12352
ID   AAU12352 standard; Protein; 243 AA.
XX
AC   AAU12352;
XX
DT   24-OCT-2001  (first entry)
XX
DE   Human PRO344 polypeptide sequence.
XX
KW   Human secretory and transmembrane; PRO; mammalian; cancer; lung;
KW   breast; prostate; cervical; tumour necrosis factor-alpha; TNF-alpha;
KW   cartilage; ear; proliferation; glucose; free fatty acid; skeletal muscle;
KW   adipocyte; A-peptide; factor VIIA; gene therapy.
XX
OS   Homo sapiens.
XX
PN   WO200140466-A2.
XX
PD   07-JUN-2001.
XX
PF   01-DEC-2000; 2000WO-US32678.
XX
PR   01-DEC-1999;   99WO-US28301.
PR   01-DEC-1999;   99WO-US28634.
PR   02-DEC-1999;   99WO-US28551.
PR   02-DEC-1999;   99WO-US28564.
PR   02-DEC-1999;   99WO-US28565.
PR   09-DEC-1999;   99US-0170262.
PR   16-DEC-1999;   99WO-US30095.
PR   20-DEC-1999;   99WO-US30911.
PR   20-DEC-1999;   99WO-US30999.
PR   30-DEC-1999;   99WO-US31243.
PR   06-JAN-2000; 2000WO-US00277.
PR   06-JAN-2000; 2000WO-US00376.
PR   11-FEB-2000; 2000WO-US03565.
PR   18-FEB-2000; 2000WO-US04341.
PR   18-FEB-2000; 2000WO-US04342.
PR   22-FEB-2000; 2000WO-US04414.
PR   24-FEB-2000; 2000WO-US04914.
PR   24-FEB-2000; 2000WO-US05004.
PR   01-MAR-2000; 2000WO-US05601.
PR   20-MAR-2000; 2000WO-US07377.
PR   21-MAR-2000; 2000WO-US07532.
PR   30-MAR-2000; 2000WO-US08439.
PR   17-MAY-2000; 2000WO-US13705.
PR   22-MAY-2000; 2000WO-US14042.
PR   30-MAY-2000; 2000WO-US14941.
PR   02-JUN-2000; 2000WO-US15264.
PR   10-NOV-2000; 2000WO-US30873.
XX
```

```
PA   (GETH ) GENENTECH INC.
XX
PI   Baker KP,  Beresini M,  Deforge L,  Desnoyers L,  Filvaroff E,  Gao W;
PI   Gerritsen ME,  Goddard A,  Godowski PJ,  Gurney AL,  Sherwood S;
PI   Smith V,  Stewart TA,  Tumas D,  Watanabe CK,  Wood WI,  Zhang Z;
XX
DR   WPI; 2001-408281/43.
DR   N-PSDB; AAS21424.
XX
PT   Isolated, secretory and transmembrane PRO polypeptide used to detect
PT   other PRO polypeptides, link bioactive molecules to cells expressing
PT   PRO polypeptides, and detect the presence of mammalian tumours e.g.
PT   lung, breast, prostate, cervical -
XX
PS   Claim 12; Fig 362; 813pp; English.
XX
CC   AAU12172-AAU12446 represent novel human secretory and transmembrane
CC   PRO polypeptides. The PRO polypeptides are useful to detect other
CC   PRO polypeptides, to link bioactive molecules to cells expressing
CC   PRO polypeptides, to modulate biological activities of cells expressing
CC   PRO polypeptides, and to detect the presence of mammalian lung, colon,
CC   breast, prostate, rectal, cervical or liver tumours by comparing PRO
CC   polypeptide expression in a cell sample to that in a control sample.
CC   Some of the 275 sequences are also useful to stimulate the release of
CC   tumour necrosis factor-alpha (TNF-alpha) from human blood, the
CC   proliferation or differentiation of chondrocytes, the proliferation or
CC   gene expression in pericyte cells, the release of proteoglycans from
CC   cartilage, the proliferation of inner ear utricular supporting cells or
CC   of T-lymphocytes, the release of a cytokine from peripheral blood
CC   monocytes (PBMCs), or the proliferation of endothelial cells. Some of
CC   the PRO polypeptides may modulate glucose or free fatty acid uptake by
CC   skeletal muscle cells or by adipocytes; or inhibit binding of A-peptide
CC   to factor VIIA. The PRO polypeptides can be used in assays to identify
CC   molecules involved in binding interactions. The polynucleotides encoding
CC   PRO polypeptides can be used to generate probes, antisense RNA/DNA,
CC   transgenic or knock out animals and can be used in gene therapy.
XX
SQ   Sequence   243 AA;
```

```
  Query Match             100.0%;  Score 243;  DB 22;  Length 243;
  Best Local Similarity   100.0%;  Pred. No. 4.9e-212;
  Matches  243;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MRPLLVLLLLGLAAGSPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MRPLLVLLLLGLAAGSPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPG 60

Qy         61 APGEKGEGGRPGLPGPRGDPGPRGEAGPAGPTGPAGECSVPPRSAFSAKRSESRVPPPSD 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 APGEKGEGGRPGLPGPRGDPGPRGEAGPAGPTGPAGECSVPPRSAFSAKRSESRVPPPSD 120

Qy        121 APLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGESIASFFQ 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 APLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGESIASFFQ 180

Qy        181 FFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWHSSP 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 FFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWHSSP 240

Qy        241 VFA 243
              |||
Db        241 VFA 243
```

```
RESULT 9
AAB65815
ID   AAB65815 standard; Protein; 243 AA.
XX
AC   AAB65815;
XX
DT   28-MAR-2001  (first entry)
XX
DE   Human TANGO 253 SEQ ID NO: 3.
XX
KW   Human; mouse; secreted protein; TANGO253; TANGO 257; TANGO 281;
KW   INTERCEPT 258; coronary disorder; olfactory disorder;
KW   neurological disorder; pulmonary disorder; immunological disorder;
KW   developmental disorder; kidney disorder.
XX
OS   Homo sapiens.
XX
PN   WO200078808-A1.
XX
PD   28-DEC-2000.
XX
PF   19-JUN-2000; 2000WO-US16883.
XX
PR   18-JUN-1999;   99US-0336536.
XX
PA   (MILL-) MILLENNIUM PHARM INC.
XX
PI   Leiby KR,  McKay C,  Bossone S;
XX
DR   WPI; 2001-050109/06.
XX
PT   New nucleic acids for treating diseases and disorders, e.g.
PT   atherosclerosis, infection, autoimmune diseases, obesity, ear
PT   disorders, brain disorders, tumors, diabetes, arthritis, multiple
PT   sclerosis and asthma -
XX
PS   Claim 9; Page 211-212; 332pp; English.
XX
CC   The present invention provides the protein and coding sequences of the
CC   human and murine secreted or transmembrane proteins TANGO 253, TANGO 257,
CC   TANGO 281 and INTERCEPT 258. These are useful in the treatment of
CC   coronary, pulmonary, olfactory, immunological, neurological,
CC   developmental and kidney disorders.
XX
SQ   Sequence   243 AA;
```

```
  Query Match             100.0%;  Score 243;  DB 22;  Length 243;
  Best Local Similarity   100.0%;  Pred. No. 4.9e-212;
  Matches  243;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MRPLLVLLLLGLAAGSPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MRPLLVLLLLGLAAGSPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPG 60

Qy         61 APGEKGEGGRPGLPGPRGDPGPRGEAGPAGPTGPAGECSVPPRSAFSAKRSESRVPPPSD 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 APGEKGEGGRPGLPGPRGDPGPRGEAGPAGPTGPAGECSVPPRSAFSAKRSESRVPPPSD 120

Qy        121 APLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGESIASFFQ 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 APLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGESIASFFQ 180

Qy        181 FFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWHSSP 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 FFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWHSSP 240

Qy        241 VFA 243
              |||
Db        241 VFA 243
```

```
RESULT 10
AAB49593
ID   AAB49593 standard; Protein; 243 AA.
XX
AC   AAB49593;
XX
DT   13-MAR-2001  (first entry)
XX
DE   Human adipocyte complement related protein homolog zsig39.
XX
KW   Human; zacrp5; gene therapy; complement inhibition; C1q domain;
```

KW adipocyte complement related protein homolog;
KW inflammation; hormone secretion; inositol phosphate; arachidonate;
KW phospholipase C activation; gastric emptying; neutrophil activation;
KW superoxide anion production; antimicrobial; acute vascular injury;
KW wound healing; zsig39.
XX
OS Homo sapiens.
XX
PN WO200073444-A1.
XX
PD 07-DEC-2000.
XX
PF 18-MAY-2000; 2000WO-US13608.
XX
PR 27-MAY-1999; 99US-0321372.
XX
PA (ZYMO ) ZYMOGENETICS INC.
XX
PI Piddington CS, Sheppard PO;
XX
DR WPI; 2001-061531/07.
XX
PT Novel adipocyte complement related protein homolog, ZACRP5, useful for
PT diagnosing and treating inflammation, vascular injury microbial
PT infections, and in wound healing -
XX
PS Disclosure; Fig 1; 121pp; English.
XX
CC The present invention relates to human adipocyte complement related
CC protein homolog, zacrp5 protein and coding sequence (see AAB49590 and
CC AAC90045). ZACRP5 has a carboxyl-terminal C1q domain. The zacrp5 gene is
CC located on human chromosome 16. zacrp5 gene and protein are useful for
CC diagnosing and treating inflammations, for determining arterial
CC remodelling, for modulating calcium ion concentration, hormone
CC secretion, DNA synthesis or cell growth, inositol phosphate turnover,
CC arachidonate release, phospholipase C activation, gastric emptying, human
CC neutrophil activation or ADCC capability and superoxide anion production.
CC zacrp5 gene and protein are also useful as antimicrobial applications,
CC preferably against bacteria and virus, for complement inhibition, for
CC treating acute vascular injury, and for wound healing. The present
CC sequence is human adipocyte complement related protein homolog, zsig39
CC protein. This protein was used in a sequence homology alignment with
CC zacrp5 protein.
XX
SQ Sequence 243 AA;

Query Match 100.0%; Score 243; DB 22; Length 243;
Best Local Similarity 100.0%; Pred. No. 4.9e-212;
Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MRPLLVLLLLGLAAGSPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MRPLLVLLLLGLAAGSPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPG 60

Qy         61 APGEKGEGGRPGLPGPRGDPGPRGEAGPAGPTGPAGECSVPPRSAFSAKRSESRVPPPSD 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 APGEKGEGGRPGLPGPRGDPGPRGEAGPAGPTGPAGECSVPPRSAFSAKRSESRVPPPSD 120

Qy        121 APLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGESIASFFQ 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 APLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGESIASFFQ 180

Qy        181 FFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWHSSP 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 FFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWHSSP 240

Qy        241 VFA 243
              |||
Db        241 VFA 243
```

RESULT 11
AAB49599

ID AAB49599 standard; Protein; 243 AA.
XX
AC AAB49599;
XX
DT 13-MAR-2001 (first entry)
XX
DE Human adipocyte complement related protein homolog zsig39.
XX
KW Human; zacrp6; gene therapy; complement inhibition; C1q domain;
KW adipocyte complement related protein homolog;
KW inflammation; hormone secretion; inositol phosphate; arachidonate;
KW phospholipase C activation; gastric emptying; neutrophil activation;
KW superoxide anion production; antimicrobial; acute vascular injury;
KW wound healing; zsig39.
XX
OS Homo sapiens.
XX
PN WO200073446-A2.
XX
PD 07-DEC-2000.
XX
PF 22-MAY-2000; 2000WO-US14024.
XX
PR 27-MAY-1999; 99US-0321262.
XX
PA (ZYMO ) ZYMOGENETICS INC.
XX
PI Piddington CS, Sheppard PO;
XX
DR WPI; 2001-061532/07.
XX
PT Novel adipocyte complement related protein homolog, ZACRP6, useful as
PT modulators of neurotransmission and for treating disseminated
PT intravascular coagulation, arteriosclerosis and acute vascular injury
PT -
XX
PS Disclosure; Fig 1; 119pp; English.
XX
CC The present invention relates to human adipocyte complement related
CC protein homolog, ZACRP6 protein and coding sequence (see AAB49596 and
CC AAC90051). ZACRP6 has a carboxyl-terminal C1q domain. The zacrp6 gene is
CC located on human chromosome 21q. zacrp6 gene and protein are useful for
CC diagnosing and treating inflammations, for determining arterial
CC remodelling, for modulating calcium ion concentration, hormone
CC secretion, DNA synthesis or cell growth, inositol phosphate turnover,
CC arachidonate release, phospholipase C activation, gastric emptying, human
CC neutrophil activation or ADCC capability and superoxide anion production.
CC zacrp6 gene and protein are also useful as antimicrobial applications,
CC preferably against bacteria and virus, for complement inhibition, for
CC treating acute vascular injury, disseminated intravascular coagulation,
CC arteriosclerosis and for wound healing. The present sequence is human
CC adipocyte complement related protein homolog zsig39. This protein was
CC used in a sequence homology comparison with ZACRP6 protein.
XX
SQ Sequence 243 AA;

Query Match 100.0%; Score 243; DB 22; Length 243;
Best Local Similarity 100.0%; Pred. No. 4.9e-212;
Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MRPLLVLLLLGLAAGSPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MRPLLVLLLLGLAAGSPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPG 60

Qy         61 APGEKGEGGRPGLPGPRGDPGPRGEAGPAGPTGPAGECSVPPRSAFSAKRSESRVPPPSD 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 APGEKGEGGRPGLPGPRGDPGPRGEAGPAGPTGPAGECSVPPRSAFSAKRSESRVPPPSD 120

Qy        121 APLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGESIASFFQ 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 APLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGESIASFFQ 180

Qy        181 FFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWHSSP 240
```

```
Db        181 FFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWHSSP 240

Qy        241 VFA 243
              |||
Db        241 VFA 243
```

RESULT 12
AAB65816
ID  AAB65816 standard; Protein; 228 AA.
XX
AC  AAB65816;
XX
DT  28-MAR-2001  (first entry)
XX
DE  Human mature TANGO 253 SEQ ID NO: 4.
XX
KW  Human; mouse; secreted protein; TANGO253; TANGO 257; TANGO 281;
KW  INTERCEPT 258; coronary disorder; olfactory disorder;
KW  neurological disorder; pulmonary disorder; immunological disorder;
KW  developmental disorder; kidney disorder.
XX
OS  Homo sapiens.
XX
PN  WO200078808-A1.
XX
PD  28-DEC-2000.
XX
PF  19-JUN-2000; 2000WO-US16883.
XX
PR  18-JUN-1999;  99US-0336536.
XX
PA  (MILL-) MILLENNIUM PHARM INC.
XX
PI  Leiby KR,  McKay C,  Bossone S;
XX
DR  WPI; 2001-050109/06.
XX
PT  New nucleic acids for treating diseases and disorders, e.g.
PT  atherosclerosis, infection, autoimmune diseases, obesity, ear
PT  disorders, brain disorders, tumors, diabetes, arthritis, multiple
PT  sclerosis and asthma -
XX
PS  Disclosure; Page 212-213; 332pp; English.
XX
CC  The present invention provides the protein and coding sequences of the
CC  human and murine secreted or transmembrane proteins TANGO 253, TANGO 257,
CC  TANGO 281 and INTERCEPT 258. These are useful in the treatment of
CC  coronary, pulmonary, olfactory, immunological, neurological,
CC  developmental and kidney disorders.
XX
SQ  Sequence  228 AA;

```
  Query Match           93.8%;  Score 228;  DB 22;  Length 228;
  Best Local Similarity  100.0%;  Pred. No. 1.8e-198;
  Matches  228;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         16 SPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPGAPGEKGEGGRPGLPG 75
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 SPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPGAPGEKGEGGRPGLPG 60

Qy         76 PRGDPGPRGEAGPAGPTGPAGECSVPPRSAFSAKRSESRVPPPSDAPLPFDRVLVNEQGH 135
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 PRGDPGPRGEAGPAGPTGPAGECSVPPRSAFSAKRSESRVPPPSDAPLPFDRVLVNEQGH 120

Qy        136 YDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGESIASFFQFFGGWPKPASLSGGA 195
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 YDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGESIASFFQFFGGWPKPASLSGGA 180

Qy        196 MVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWHSSPVFA 243
              ||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 MVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWHSSPVFA 228
```

RESULT 13
ABG12724
ID  ABG12724 standard; Protein; 220 AA.
XX
AC  ABG12724;
XX
DT  18-FEB-2002  (first entry)
XX
DE  Novel human diagnostic protein #12715.
XX
KW  Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW  food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS  Homo sapiens.
XX
PN  WO200175067-A2.
XX
PD  11-OCT-2001.
XX
PF  30-MAR-2001; 2001WO-US08631.
XX
PR  31-MAR-2000; 2000US-0540217.
PR  23-AUG-2000; 2000US-0649167.
XX
PA  (HYSE-) HYSEQ INC.
XX
PI  Drmanac RT,  Liu C,  Tang YT;
XX
DR  WPI; 2001-639362/73.
DR  N-PSDB; AAS76911.
XX
PT  New isolated polynucleotide and encoded polypeptides, useful in
PT  diagnostics, forensics, gene mapping, identification of mutations
PT  responsible for genetic disorders or other traits and to assess
PT  biodiversity -
XX
PS  Claim 20; SEQ ID No 43083; 103pp; English.
XX
CC  The invention relates to isolated polynucleotide (I) and
CC  polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC  polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC  and gene mapping, and in recombinant production of (II). The
CC  polynucleotides are also used in diagnostics as expressed sequence tags
CC  for identifying expressed genes. (I) is useful in gene therapy techniques
CC  to restore normal activity of (II) or to treat disease states involving
CC  (II). (II) is useful for generating antibodies against it, detecting or
CC  quantitating a polypeptide in tissue, as molecular weight markers and as
CC  a food supplement. (II) and its binding partners are useful in medical
CC  imaging of sites expressing (II). (I) and (II) are useful for treating
CC  disorders involving aberrant protein expression or biological activity.
CC  The polypeptide and polynucleotide sequences have applications in
CC  diagnostics, forensics, gene mapping, identification of mutations
CC  responsible for genetic disorders or other traits to assess biodiversity
CC  and to produce other types of data and products dependent on DNA and
CC  amino acid sequences. ABG00010-ABG30377 represent novel human
CC  diagnostic amino acid sequences of the invention.
CC  Note: The sequence data for this patent did not appear in the printed
CC  specification, but was obtained in electronic format directly from WIPO
CC  at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ  Sequence  220 AA;

```
  Query Match           87.7%;  Score 213;  DB 22;  Length 220;
  Best Local Similarity  100.0%;  Pred. No. 6.9e-185;
  Matches  213;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         25 PSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPGAPGEKGEGGRPGLPGPRGDPGPRG 84
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 PSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPGAPGEKGEGGRPGLPGPRGDPGPRG 60

Qy         85 EAGPAGPTGPAGECSVPPRSAFSAKRSESRVPPPSDAPLPFDRVLVNEQGHYDAVTGKFT 144
```

```
Db          61 EAGPAGPTGPAGECSVPPRSAFSAKRSESRVPPPSDAPLPFDRVLVNEQGHYDAVTGKFT 120

Qy         145 CQVPGVYYFAVHATVYRASLQFDLVKNGESIASFFQFFGGWPKPASLSGGAMVRLEPEDQ 204

Db         121 CQVPGVYYFAVHATVYRASLQFDLVKNGESIASFFQFFGGWPKPASLSGGAMVRLEPEDQ 180

Qy         205 VWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWH 237

Db         181 VWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWH 213
```

RESULT 14
AAB65888
ID AAB65888 standard; Protein; 243 AA.
XX
AC AAB65888;
XX
DT 28-MAR-2001 (first entry)
XX
DE Human secreted protein related protein SEQ ID NO: 102.
XX
KW Human; mouse; secreted protein; TANGO253; TANGO 257; TANGO 281;
KW INTERCEPT 258; coronary disorder; olfactory disorder;
KW neurological disorder; pulmonary disorder; immunological disorder;
KW developmental disorder; kidney disorder.
XX
OS Homo sapiens.
XX
PN WO200078808-A1.
XX
PD 28-DEC-2000.
XX
PF 19-JUN-2000; 2000WO-US16883.
XX
PR 18-JUN-1999; 99US-0336536.
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
PI Leiby KR, McKay C, Bossone S;
XX
DR WPI; 2001-050109/06.
XX
PT New nucleic acids for treating diseases and disorders, e.g.
PT atherosclerosis, infection, autoimmune diseases, obesity, ear
PT disorders, brain disorders, tumors, diabetes, arthritis, multiple
PT sclerosis and asthma -
XX
PS Disclosure; Page 270-271; 332pp; English.
XX
CC The present invention provides the protein and coding sequences of the
CC human and murine secreted or transmembrane proteins TANGO 253, TANGO 257,
CC TANGO 281 and INTERCEPT 258. These are useful in the treatment of
CC coronary, pulmonary, olfactory, immunological, neurological,
CC developmental and kidney disorders.
XX
SQ Sequence 243 AA;

Query Match 76.1%; Score 185; DB 22; Length 243;
Best Local Similarity 100.0%; Pred. No. 1.8e-159;
Matches 185; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          59 PGAPGEKGEGGRPGLPGPRGDPGPRGEAGPAGPTGPAGECSVPPRSAFSAKRSESRVPPP 118

Db          59 PGAPGEKGEGGRPGLPGPRGDPGPRGEAGPAGPTGPAGECSVPPRSAFSAKRSESRVPPP 118

Qy         119 SDAPLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGESIASF 178

Db         119 SDAPLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGESIASF 178

Qy         179 FQFFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWHS 238

Db         179 FQFFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWHS 238

Qy         239 SPVFA 243

Db         239 SPVFA 243
```

RESULT 15
AAB65891
ID AAB65891 standard; Protein; 243 AA.
XX
AC AAB65891;
XX
DT 28-MAR-2001 (first entry)
XX
DE Human secreted protein related protein SEQ ID NO: 108.
XX
KW Human; mouse; secreted protein; TANGO253; TANGO 257; TANGO 281;
KW INTERCEPT 258; coronary disorder; olfactory disorder;
KW neurological disorder; pulmonary disorder; immunological disorder;
KW developmental disorder; kidney disorder.
XX
OS Homo sapiens.
XX
PN WO200078808-A1.
XX
PD 28-DEC-2000.
XX
PF 19-JUN-2000; 2000WO-US16883.
XX
PR 18-JUN-1999; 99US-0336536.
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
PI Leiby KR, McKay C, Bossone S;
XX
DR WPI; 2001-050109/06.
XX
PT New nucleic acids for treating diseases and disorders, e.g.
PT atherosclerosis, infection, autoimmune diseases, obesity, ear
PT disorders, brain disorders, tumors, diabetes, arthritis, multiple
PT sclerosis and asthma -
XX
PS Disclosure; Page 274; 332pp; English.
XX
CC The present invention provides the protein and coding sequences of the
CC human and murine secreted or transmembrane proteins TANGO 253, TANGO 257,
CC TANGO 281 and INTERCEPT 258. These are useful in the treatment of
CC coronary, pulmonary, olfactory, immunological, neurological,
CC developmental and kidney disorders.
XX
SQ Sequence 243 AA;

Query Match 71.6%; Score 174; DB 22; Length 243;
Best Local Similarity 100.0%; Pred. No. 1.7e-149;
Matches 174; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy           1 MRPLLVLLLLGLAAGSPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPG 60

Db           1 MRPLLVLLLLGLAAGSPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPG 60

Qy          61 APGEKGEGGRPGLPGPRGDPGPRGEAGPAGPTGPAGECSVPPRSAFSAKRSESRVPPPSD 120

Db          61 APGEKGEGGRPGLPGPRGDPGPRGEAGPAGPTGPAGECSVPPRSAFSAKRSESRVPPPSD 120

Qy         121 APLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGES 174

Db         121 APLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGES 174
```

Search completed: June 18, 2003, 15:03:51
Job time : 77 secs

GenCore version 5.1.6


OM protein - protein search, using sw model

Run on: June 18, 2003, 15:01:43 ; Search time 40 Seconds
(without alignments)
584.017 Million cell updates/sec

Title: US-09-943-851A-42
Perfect score: 243
Sequence: 1 MRPLLVLLLLGLAAGSPPLD.....DSTFSGFLVYSDWHSSPVFA 243

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 283224 seqs, 96134422 residues

Word size : 0

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : PIR_73:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 219 | 90.1 | 219 | 2 | T14782 | hypothetical prote |
| 2 | 11 | 4.5 | 289 | 2 | T34241 | hypothetical prote |
| 3 | 11 | 4.5 | 369 | 2 | S33603 | surfactant protein |
| 4 | 11 | 4.5 | 374 | 1 | A42046 | surfactant protein |
| 5 | 11 | 4.5 | 375 | 1 | A45225 | pulmonary surfacta |
| 6 | 11 | 4.5 | 1670 | 1 | CGHU3B | collagen alpha 3(I |
| 7 | 11 | 4.5 | 2944 | 2 | A54849 | collagen alpha 1(V |
| 8 | 10 | 4.1 | 210 | 2 | B44984 | collagen - nematod |
| 9 | 10 | 4.1 | 247 | 1 | LNRBPS | pulmonary surfacta |
| 10 | 10 | 4.1 | 248 | 1 | LNHUP6 | pulmonary surfacta |
| 11 | 10 | 4.1 | 248 | 1 | LNHUP1 | pulmonary surfacta |
| 12 | 10 | 4.1 | 248 | 1 | LNDGPS | pulmonary surfacta |
| 13 | 10 | 4.1 | 248 | 1 | LNRTPS | pulmonary surfacta |
| 14 | 10 | 4.1 | 248 | 2 | I51921 | pulmonary surfacta |
| 15 | 10 | 4.1 | 248 | 2 | A48853 | pulmonary surfacta |
| 16 | 10 | 4.1 | 447 | 2 | T42987 | hypothetical prote |
| 17 | 10 | 4.1 | 547 | 2 | A36046 | collagen alpha cha |
| 18 | 10 | 4.1 | 694 | 2 | H95012 | hypothetical prote |
| 19 | 10 | 4.1 | 702 | 2 | C97884 | hypothetical prote |
| 20 | 10 | 4.1 | 812 | 2 | S31521 | collagen COLF1 - f |
| 21 | 10 | 4.1 | 1049 | 1 | CGBO7S | collagen alpha 1(I |
| 22 | 10 | 4.1 | 1464 | 2 | S59856 | collagen alpha 1(I |
| 23 | 10 | 4.1 | 1466 | 1 | CGHU7L | collagen alpha 1(I |
| 24 | 10 | 4.1 | 1761 | 2 | T13990 | collagen type IV a |
| 25 | 10 | 4.1 | 2551 | 2 | B98047 | hypothetical prote |
| 26 | 9 | 3.7 | 126 | 2 | S53786 | collagen alpha cha |
| 27 | 9 | 3.7 | 246 | 2 | S38344 | CDEI-binding prote |
| 28 | 9 | 3.7 | 265 | 2 | G82013 | capsule polysaccha |
| 29 | 9 | 3.7 | 356 | 2 | S16907 | collagen alpha 1(I |
| 30 | 9 | 3.7 | 636 | 2 | S41067 | collagen alpha 1(I |
| 31 | 9 | 3.7 | 751 | 2 | A49974 | beta-amyloid precu |
| 32 | 9 | 3.7 | 765 | 2 | S42880 | amyloid precursor- |
| 33 | 9 | 3.7 | 964 | 1 | CGCH2S | collagen alpha 2(I |
| 34 | 9 | 3.7 | 1402 | 2 | I46707 | translation initia |
| 35 | 9 | 3.7 | 1669 | 1 | CGMS4B | collagen alpha 1(I |
| 36 | 9 | 3.7 | 1752 | 2 | A45407 | collagen alpha 3(I |
| 37 | 9 | 3.7 | 2288 | 2 | T29999 | hypothetical prote |
| 38 | 8 | 3.3 | 46 | 2 | I65268 | collagen alpha 1(I |
| 39 | 8 | 3.3 | 52 | 2 | C83241 | hypothetical prote |
| 40 | 8 | 3.3 | 87 | 2 | H31482 | hypothetical prote |
| 41 | 8 | 3.3 | 149 | 2 | B82017 | probable integral |
| 42 | 8 | 3.3 | 149 | 2 | B81002 | conserved hypothet |
| 43 | 8 | 3.3 | 164 | 2 | S76920 | hypothetical prote |
| 44 | 8 | 3.3 | 170 | 2 | S74248 | antibacterial pept |
| 45 | 8 | 3.3 | 170 | 2 | I38932 | CAP18 precursor - |

ALIGNMENTS

RESULT 1
T14782
hypothetical protein DKFZp586B0621.1 - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 02-Sep-2000
C;Accession: T14782
R;Ottenwaelder, B.; Obermaier, B.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, August 1999
A;Reference number: Z18184
A;Accession: T14782
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-219 <OTT>
A;Cross-references: EMBL:AL110261
A;Experimental source: adult uterus; clone DKFZp586B0621
C;Genetics:
A;Note: DKFZp586B0621.1
C;Superfamily: complement C1q carboxyl-terminal homology

Query Match 90.1%; Score 219; DB 2; Length 219;
Best Local Similarity 100.0%; Pred. No. 1.5e-199;
Matches 219; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          25 PSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPGAPGEKGEGGRPGLPGPRGDPGPRG 84
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           1 PSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPGAPGEKGEGGRPGLPGPRGDPGPRG 60

Qy          85 EAGPAGPTGPAGECSVPPRSAFSAKRSESRVPPPSDAPLPFDRVLVNEQGHYDAVTGKFT 144
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          61 EAGPAGPTGPAGECSVPPRSAFSAKRSESRVPPPSDAPLPFDRVLVNEQGHYDAVTGKFT 120

Qy         145 CQVPGVYYFAVHATVYRASLQFDLVKNGESIASFFQFFGGWPKPASLSGGAMVRLEPEDQ 204
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         121 CQVPGVYYFAVHATVYRASLQFDLVKNGESIASFFQFFGGWPKPASLSGGAMVRLEPEDQ 180

Qy         205 VWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWHSSPVFA 243
               |||||||||||||||||||||||||||||||||||||||
Db         181 VWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWHSSPVFA 219
```

RESULT 2
T34241
hypothetical protein F26F12.1 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 21-Jan-2000
C;Accession: T34241
R;Wilson, R.; Bentley, D.; Gattung, S.
submitted to the EMBL Data Library, April 1996
A;Description: The sequence of C. elegans cosmid F26F12.
A;Reference number: Z21493
A;Accession: T34241
A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: DNA
A;Residues: 1-289 <WIL>
A;Cross-references: EMBL:U55373; PIDN:AAC25888.1; GSPDB:GN00023; CESP:F26F12.1
A;Experimental source: strain Bristol N2; clone F26F12
C;Genetics:
A;Gene: CESP:F26F12.1
A;Map position: 5
A;Introns: 45/3
C;Superfamily: unassigned collagens

Query Match 4.5%; Score 11; DB 2; Length 289;
Best Local Similarity 100.0%; Pred. No. 0.016;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 80 PGPRGEAGPAG 90
||||||||||||
Db 222 PGPRGEAGPAG 232

RESULT 3
S33603
surfactant protein D - bovine
C;Species: Bos primigenius taurus (cattle)
C;Date: 02-Dec-1993 #sequence_revision 01-Sep-1995 #text_change 17-Mar-1999
C;Accession: S33603
R;Lim, B.L.; Lu, J.; Reid, K.B.M.
Immunology 78, 159-165, 1993
A;Title: Structural similarity between bovine conglutinin and bovine lung surfactant pro
A;Reference number: S33603; MUID:93170856; PMID:8436402
A;Accession: S33603
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-369 <LIM>
C;Superfamily: pulmonary surfactant protein D; C-type lectin homology
F;248-367/Domain: C-type lectin homology <LCH>

Query Match 4.5%; Score 11; DB 2; Length 369;
Best Local Similarity 100.0%; Pred. No. 0.02;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 45 GLPGRDGRDGR 55
||||||||||||
Db 46 GLPGRDGRDGR 56

RESULT 4
A42046
surfactant protein D - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C;Accession: A42046
R;Shimizu, H.; Fisher, J.H.; Papst, P.; Benson, B.; Lau, K.; Mason, R.J.; Voelker, D.R.
J. Biol. Chem. 267, 1853-1857, 1992
A;Title: Primary structure of rat pulmonary surfactant protein D. cDNA and deduced amino
A;Reference number: A42046; MUID:92112913; PMID:1370483
A;Accession: A42046
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-374 <SHI>
A;Cross-references: GB:M81231; NID:g207035; PIDN:AAA42170.1; PID:g207036
A;Experimental source: lung
A;Note: sequence extracted from NCBI backbone (NCBIN:76027, NCBIP:76031)
C;Superfamily: pulmonary surfactant protein D; C-type lectin homology
F;253-372/Domain: C-type lectin homology <LCH>

Query Match 4.5%; Score 11; DB 1; Length 374;
Best Local Similarity 100.0%; Pred. No. 0.02;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 45 GLPGRDGRDGR 55
||||||||||||
Db 45 GLPGRDGRDGR 55

RESULT 5
A45225
pulmonary surfactant protein D precursor - human
N;Alternate names: SP-D
C;Species: Homo sapiens (man)
C;Date: 16-Apr-1999 #sequence_revision 16-Apr-1999 #text_change 22-Jun-1999
C;Accession: A45225; S23434; S24555; S44420; S18382; A56776
R;Crouch, E.; Rust, K.; Veile, R.; Donis-Keller, H.; Grosso, L.
J. Biol. Chem. 268, 2976-2983, 1993
A;Title: Genomic organization of human surfactant protein D (SP-D). SP-D is encoded o
A;Reference number: A45225; MUID:93155122; PMID:8428971
A;Accession: A45225
A;Molecule type: DNA
A;Residues: 1-375 <CRO>
A;Cross-references: GB:L05483; GB:L05484; GB:L05485; NID:g292505; PIDN:AAB59450.1; PI
A;Experimental source: placenta
A;Note: sequence extracted from NCBI backbone (NCBIP:124316)
R;Lu, J.; Willis, A.C.; Reid, K.B.M.
Biochem. J. 284, 795-802, 1992
A;Title: Purification, characterization and cDNA cloning of human lung surfactant pro
A;Reference number: S23434; MUID:92322003; PMID:1339284
A;Accession: S23434
A;Molecule type: mRNA
A;Residues: 1-30,'T',32-121,'P',123-179,'A',181-375 <LUJ1>
A;Cross-references: EMBL:X65018; NID:g34766; PIDN:CAA46152.1; PID:g34767
A;Experimental source: lung
A;Accession: S24555
A;Molecule type: protein
A;Residues: 214-234,'X',236,'XX',239-241 <LUJ2>
R;Hoppe, H.J.; Barlow, P.N.; Reid, K.B.M.
FEBS Lett. 344, 191-195, 1994
A;Title: A parallel three stranded alpha-helical bundle at the nucleation site of col
A;Reference number: S44420; MUID:94244769; PMID:8187882
A;Accession: S44420
A;Molecule type: mRNA
A;Residues: 202-257 <HOP>
R;Rust, K.; Grosso, L.; Zhang, V.; Chang, D.; Persson, A.; Longmore, W.; Cai, G.Z.; C
Arch. Biochem. Biophys. 290, 116-126, 1991
A;Title: Human surfactant protein D: SP-D contains a C-type lectin carbohydrate recog
A;Reference number: S18382; MUID:91378578; PMID:1898081
A;Accession: S18382
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 'F',60-205,'P',207-374,'HF' <RUS>
A;Cross-references: GB:L05485; NID:g292505
A;Note: corrections to this sequence are reported in reference A56776
R;Crouch, E.; Persson, A.; Chang, D.
Am. J. Pathol. 142, 241-248, 1993
A;Title: Accumulation of surfactant protein D in human pulmonary alveolar proteinosis
A;Reference number: A56776; MUID:93142849; PMID:8424457
A;Accession: A56776
A;Status: preliminary
A;Molecule type: protein
A;Residues: 46-58,'F',60-62,'E',64-72;223-227,'X',229-239,'P',241-245,'X',247-256,'X'
A;Cross-references: PIDN:AAB25037.1; PID:g263973; PIDN:AAB25038.1; PID:g263974
A;Experimental source: bronchoalveolar lavage
A;Note: sequence extracted from NCBI backbone (NCBIP:123024, NCBIP:123023); sequence
C;Comment: Pulmonary surfactant is a complex of phospholipids and proteins that lower
C;Comment: This protein is synthesized by alveolar type II cells.
C;Genetics:
A;Gene: GDB:SFTPD; SFTP4; SP-D
A;Cross-references: GDB:132674; OMIM:178635
A;Map position: 10q22.2-10q23.1
C;Superfamily: pulmonary surfactant protein D; C-type lectin homology
C;Keywords: blocked amino end; calcium; glycoprotein; hydroxylysine; hydroxyproline;
F;1-20/Domain: signal sequence #status predicted <SIG>
F;21-375/Product: pulmonary surfactant protein D #status predicted <MAT>
F;21-45/Domain: non-collagenous #status predicted <NC1>
F;46-222/Domain: collagenous #status predicted <COL>
F;223-375/Domain: non-collagenous #status predicted <NC2>
F;254-373/Domain: C-type lectin homology <LCH>
F;90/Binding site: carbohydrate (Asn) (covalent) #status predicted

F;281-373,351-365/Disulfide bonds: #status predicted

Query Match 4.5%; Score 11; DB 1; Length 375;
Best Local Similarity 100.0%; Pred. No. 0.021;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 45 GLPGRDGRDGR 55
||||||||||||
Db 46 GLPGRDGRDGR 56

RESULT 6
CGHU3B
collagen alpha 3(IV) chain precursor, long splice form - human
N;Alternate names: Goodpasture antigen; procollagen alpha 3(IV) chain long splice form
C;Species: Homo sapiens (man)
C;Date: 28-Oct-1994 #sequence_revision 03-Oct-1995 #text_change 22-Jun-1999
C;Accession: A54763; A43928; A44043; A45971; A39786
R;Mariyama, M.; Leinonen, A.; Mochizuki, T.; Tryggvason, K.; Reeders, S.T.
J. Biol. Chem. 269, 23013-23017, 1994
A;Title: Complete primary structure of the human alpha3(IV) collagen chain. Coexpression
A;Reference number: A54763; MUID:94364994; PMID:8083201
A;Accession: A54763
A;Molecule type: mRNA
A;Residues: 1-1670 <MAR>
A;Cross-references: GB:X80031; NID:g577563; PID:g577564
A;Experimental source: kidney
R;Turner, N.; Mason, P.J.; Brown, R.; Fox, M.; Povey, S.; Rees, A.; Pusey, C.D.
J. Clin. Invest. 89, 592-601, 1992
A;Title: Molecular cloning of the human Goodpasture antigen demonstrates it to be the al
A;Reference number: A43928; MUID:92147878; PMID:1737849
A;Accession: A43928
A;Molecule type: mRNA
A;Residues: 1331-1524,'I',1526-1670 <TUR>
A;Cross-references: GB:M81379
A;Experimental source: kidney
R;Quinones, S.; Bernal, D.; Garcia-Sogo, M.; Elena, S.F.; Saus, J.
J. Biol. Chem. 267, 19780-19784, 1992
A;Title: Exon/intron structure of the human alpha 3(IV) gene encompassing the Goodpastur
ction.
A;Reference number: A44043; MUID:93015826; PMID:1400291
A;Accession: A44043
A;Molecule type: DNA; mRNA
A;Residues: 1386-1670 <QUI>
A;Cross-references: GB:M92993; NID:g177895; PIDN:AAA21610.1; PID:g177896
A;Note: sequence extracted from NCBI backbone (NCBIP:115597)
R;Quinones, S.; Bernal, D.; Garcia-Sogo, M.; Elena, S.F.; Saus, J.
J. Biol. Chem. 269, 17358, 1994
A;Reference number: A44738; MUID:94274734; PMID:8006044
A;Contents: annotation; erratum; correction to intronic sequence in A44043
R;Bernal, D.; Quinones, S.; Saus, J.
J. Biol. Chem. 268, 12090-12094, 1993
A;Title: The human mRNA encoding the Goodpasture antigen is alternatively spliced.
A;Reference number: A45971; MUID:93280184; PMID:8505332
A;Accession: A45971
A;Status: nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1427-1444 <BER>
A;Note: sequence extracted from NCBI backbone (NCBIP:133363); sequence incorrectly ident
R;Morrison, K.E.; Mariyama, M.; Yang-Feng, T.L.; Reeders, S.T.
Am. J. Hum. Genet. 49, 545-554, 1991
A;Title: Sequence and localization of a partial cDNA encoding the human alpha3 chain of
A;Reference number: A39786; MUID:91353570; PMID:1882840
A;Accession: A39786
A;Molecule type: mRNA
A;Residues: 1453-1593,'A',1595-1670 <MOR>
A;Cross-references: GB:S55790; NID:g234418; PIDN:AAB19637.1; PID:g234419
C;Comment: Prolines and lysines at the third position of the tripeptide repeating unit (
ed and subsequently O-glycosylated.
C;Comment: In Goodpasture's syndrome, an autoimmune response develops against an epitope
C;Genetics:
A;Gene: GDB:COL4A3
A;Cross-references: GDB:128351; OMIM:120070
A;Map position: 2q36-2q37
A;Introns: 1385/1; 1418/1; 1488/1; 1547/2; 1585/3; 1643/2 #status incomplete
A;Note: the alpha 3(IV) and alpha 4(IV) chain genes are encoded on opposite strands w
C;Complex: This minor type IV collagen is thought to form a heterotrimer of two alpha
mong trimer amino-terminal domains (with disulfide and desmosine cross-links), dimeri
er associations in the interrupted helical domain (with disulfide and desmosine cross
C;Function:
A;Description: minor structural component of extracellular basement membrane in kidne
C;Superfamily: collagen alpha 1(IV) chain
C;Keywords: alternative splicing; basement membrane; cell binding; coiled coil; extra
F;1-28/Domain: signal sequence #status predicted <SIG>
F;29-1670/Product: collagen alpha 3(IV) chain, long splice form #status predicted <MA
F;29-42/Domain: amino-terminal nonhelical, NH1 <NH1>
F;43-1438/Region: interrupted helical
F;791-793/Region: cell attachment (R-G-D) motif
F;996-998/Region: cell attachment (R-G-D) motif
F;1154-1156/Region: cell attachment (R-G-D) motif
F;1306-1308/Region: cell attachment (R-G-D) motif
F;1345-1347/Region: cell attachment (R-G-D) motif
F;1432-1434/Region: cell attachment (R-G-D) motif
F;1439-1670/Domain: carboxyl-terminal nonhelical, NC1 <NC1>
F;1451-1551/Domain: collagen IV carboxyl-terminal repeat <CT1>
F;1561-1665/Domain: collagen IV carboxyl-terminal repeat <CT2>
F;31,33,39,41,125,422,476,479,682,722,809,1387/Disulfide bonds: interchain #status pr
F;253/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;1460-1548,1493-1551/Disulfide bonds: (or 1460-1551, 1493-1548) #status predicted
F;1505-1511,1616-1622/Disulfide bonds: #status predicted
F;1570-1662,1604-1665/Disulfide bonds: (or 1570-1665, 1604-1662) #status predicted

Query Match 4.5%; Score 11; DB 1; Length 1670;
Best Local Similarity 100.0%; Pred. No. 0.074;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 71 PGLPGPRGDPG 81
||||||||||||
Db 1300 PGLPGPRGDPG 1310

RESULT 7
A54849
collagen alpha 1(VII) chain precursor - human
N;Alternate names: procollagen alpha 1(VII) chain
C;Species: Homo sapiens (man)
C;Date: 04-Nov-1994 #sequence_revision 04-Nov-1994 #text_change 20-Sep-1999
C;Accession: A54849; PH0844; S16316; I56328; A30296; I84686
R;Christiano, A.M.; Greenspan, D.S.; Lee, S.; Uitto, J.
J. Biol. Chem. 269, 20256-20262, 1994
A;Title: Cloning of human type VII collagen. Complete primary sequence of the alpha1(
A;Reference number: A54849; MUID:94327588; PMID:8051117
A;Accession: A54849
A;Status: not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-2944 <CHR>
A;Cross-references: GB:L02870; NID:g987124; PIDN:AAA75438.1; PID:g987125
R;Tanaka, T.; Takahashi, K.; Furukawa, F.; Imamura, S.
Biochem. Biophys. Res. Commun. 183, 958-963, 1992
A;Title: Molecular cloning and characterization of type VII collagen cDNA.
A;Reference number: PH0844; MUID:92231902; PMID:1567409
A;Accession: PH0844
A;Molecule type: mRNA
A;Residues: 'EFR',340-475,'RALSTASHSTLCWRATRWHPCNRGSHWTRAACEPCNRPASHRAARAG',524-528,'
A;Cross-references: DDBJ:D11152; DDBJ:D13694; NID:g453698; PIDN:BAA02853.1; PID:g4536
A;Experimental source: keratinocyte
A;Note: the authors translated the codon ACC for residues 394 and 397 as Tyr
R;Parente, M.G.; Chung, L.C.; Ryynaenen, J.; Woodley, D.T.; Wynn, K.C.; Bauer, E.A.;
Proc. Natl. Acad. Sci. U.S.A. 88, 6931-6935, 1991
A;Title: Human type VII collagen: cDNA cloning and chromosomal mapping of the gene.
A;Reference number: S16316; MUID:91334380; PMID:1871109
A;Accession: S16316
A;Molecule type: mRNA
A;Residues: 815-892,'E',894-1439 <PAR>
A;Cross-references: GB:M65158; GB:S49017; NID:g180914; PIDN:AAA96439.1; PID:g180915
A;Experimental source: keratinocyte

R;Gammon, W.R.; Abernethy, M.L.; Padilla, K.M.; Prisayanh, P.S.; Cook, M.E.; Wright, J.;
J. Invest. Dermatol. 99, 691-696, 1992
A;Title: Noncollagenous (NC1) domain of collagen VII resembles multidomain adhesion prot
A;Reference number: I56328; MUID:93107742; PMID:1469284
A;Accession: I56328
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 'EFR',372-517,'DV',520-540,'W',542-1255 <RES>
A;Cross-references: GB:S51236; NID:g262308; PIDN:AAB24637.1; PID:g262309
R;Seltzer, J.L.; Eisen, A.Z.; Bauer, E.A.; Morris, N.P.; Glanville, R.W.; Burgeson, R.E.
J. Biol. Chem. 264, 3822-3826, 1989
A;Title: Cleavage of type VII collagen by interstitial collagenase and type IV collagena
A;Reference number: A30296; MUID:89139437; PMID:2537292
A;Accession: A30296
A;Molecule type: protein
A;Residues: 'A',1240-1246,'G',1248-1250,'XE',1253-1255,'Q',1257,'E';2032,'C',2034-2041;'
A;Note: two reported peptides cannot be reliably located
R;Greenspan, D.S.
Hum. Mol. Genet. 2, 273-278, 1993
A;Title: The carboxyl-terminal half of type VII collagen, including the non-collagenous
A;Reference number: I48103; MUID:93271985; PMID:8499916
A;Accession: I84686
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 2395-2871,'S',2873-2944 <RE2>
A;Cross-references: GB:L06862; NID:g388713; PIDN:AAA89196.1; PID:g388714
R;Christiano, A.M.; Ryynaenen, M.; Uitto, J.
Proc. Natl. Acad. Sci. U.S.A. 91, 3549-3553, 1994
A;Title: Dominant dystrophic epidermolysis bullosa: identification of a Gly --> Ser subs
A;Reference number: A55255; MUID:94224777; PMID:8170945
A;Contents: annotation
C;Comment: Prolines and lysines at the third position of the tripeptide repeating unit (
ed and subsequently O-glycosylated.
C;Genetics:
A;Gene: GDB:COL7A1; EBR1; EBD1; EB
A;Cross-references: GDB:128750; OMIM:120120
A;Map position: 3p21.3-3p21.3
A;Note: defects in this gene can result in dominant and recessive dystrophic epidermolys
A;Note: there are 118 introns
C;Complex: type VII collagen is probably a homotrimer
C;Function:
A;Description: structural component of extracellular polymer associated with anchoring f
C;Superfamily: unassigned collagens; animal Kunitz-type proteinase inhibitor homology; f
C;Keywords: coiled coil; extracellular matrix; glycoprotein; hydroxylysine; hydroxyproli
F;1-16/Domain: signal sequence #status predicted <SIG>
F;17-2944/Product: collagen alpha 1(VII) chain #status predicted <MAT>
F;17-1253/Domain: amino-terminal nonhelical #status predicted <NC1>
F;36-201/Domain: von Willebrand factor type A repeat homology <VWA1>
F;231-318/Domain: fibronectin type III repeat homology <FN1>
F;327-413/Domain: fibronectin type III repeat homology <FN2>
F;414-502/Domain: fibronectin type III repeat homology <FN3>
F;508-593/Domain: fibronectin type III repeat homology <FN4>
F;598-683/Domain: fibronectin type III repeat homology <FN5>
F;686-771/Domain: fibronectin type III repeat homology <FN6>
F;776-862/Domain: fibronectin type III repeat homology <FN7>
F;864-952/Domain: fibronectin type III repeat homology <FN8>
F;954-1045/Domain: fibronectin type III repeat homology <FN9>
F;1052-1219/Domain: von Willebrand factor type A repeat homology <VWA2>
F;1170-1172/Region: cell attachment (R-G-D) motif
F;1189-1253/Region: cysteine/proline-rich
F;1254-2783/Region: interrupted helical
F;1334-1336/Region: cell attachment (R-G-D) motif
F;2008-2010/Region: cell attachment (R-G-D) motif
F;2553-2555/Region: cell attachment (R-G-D) motif
F;2784-2944/Domain: carboxyl-terminal nonhelical #status predicted <NC2>
F;2876-2929/Domain: animal Kunitz-type proteinase inhibitor homology <BPI>
F;337,786,1109/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;2167,2176,2185,2188,2664,2667,2673/Modified site: 4-hydroxyproline (Pro) #status exper
F;2625,2631/Modified site: 5-hydroxylysine (Lys) #status experimental
F;2625,2631/Binding site: carbohydrate (Lys) (covalent) #status experimental
F;2634,2802,2804/Disulfide bonds: interchain #status predicted

Query Match 4.5%; Score 11; DB 2; Length 2944;
Best Local Similarity 100.0%; Pred. No. 0.12;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 71 PGLPGPRGDPG 81
||||||||||
Db 1328 PGLPGPRGDPG 1338

RESULT 8
B44984
collagen - nematode (Haemonchus contortus) (fragment)
C;Species: Haemonchus contortus
C;Date: 28-Apr-1993 #sequence_revision 28-Apr-1993 #text_change 07-May-1999
C;Accession: B44984
R;Shamansky, L.M.; Pratt, D.; Boisvenue, R.J.; Cox, G.N.
Mol. Biochem. Parasitol. 37, 73-86, 1989
A;Title: Cuticle collagen genes of Haemonchus contortus and Caenorhabditis elegans ar
A;Reference number: A44984; MUID:90136718; PMID:2615789
A;Accession: B44984
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-210 <SHA>
A;Cross-references: GB:J04671; GB:J04670
C;Superfamily: unassigned collagens

Query Match 4.1%; Score 10; DB 2; Length 210;
Best Local Similarity 100.0%; Pred. No. 0.11;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 57 GAPGAPGEKG 66
||||||||||
Db 151 GAPGAPGEKG 160

RESULT 9
LNRBPS
pulmonary surfactant protein A precursor - rabbit
N;Alternate names: pulmonary surfactant 32K apoprotein; pulmonary surfactant-associat
C;Species: Oryctolagus cuniculus (domestic rabbit)
C;Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 16-Jul-1999
C;Accession: A29931
R;Boggaram, V.; Qing, K.; Mendelson, C.R.
J. Biol. Chem. 263, 2939-2947, 1988
A;Title: The major apoprotein of rabbit pulmonary surfactant. Elucidation of primary
A;Reference number: A29931; MUID:88139348; PMID:2830270
A;Accession: A29931
A;Molecule type: mRNA
A;Residues: 1-247 <BOG>
A;Cross-references: GB:J03542; NID:g165705; PIDN:AAA31465.1; PID:g165706
A;Note: 12-Pro was also found
A;Note: two species of mRNA, which appear to be transcribed from a single gene, could
A;Note: the amino end of the mature protein is blocked
C;Comment: Pulmonary surfactant is a complex of phospholipids and proteins that lower
C;Comment: This protein is a sialoglycoprotein synthesized by alveolar type II cells.
pendent on the presence of calcium ions.
C;Superfamily: mannose-binding lectin; C-type lectin homology
C;Keywords: acetylated amino end; alveolar proteinosis; calcium; gaseous exchange; gl
F;1-15/Domain: signal sequence #status predicted <SIG>
F;16-247/Product: pulmonary surfactant protein A #status predicted <MAT>
F;27-99/Region: collagen-like
F;126-245/Domain: C-type lectin homology <LCH>
F;16/Modified site: acetylated amino end (Ser) (in mature form) #status predicted
F;206/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 4.1%; Score 10; DB 1; Length 247;
Best Local Similarity 100.0%; Pred. No. 0.13;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 45 GLPGRDGRDG 54
||||||||||
Db 39 GLPGRDGRDG 48

RESULT 10
LNHUP6
pulmonary surfactant protein A precursor (clone 6A) - human
N;Alternate names: pulmonary surfactant 32K apoprotein; pulmonary surfactant-associated
C;Species: Homo sapiens (man)
C;Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 16-Jul-1999
C;Accession: A25720
R;Floros, J.; Steinbrink, R.; Jacobs, K.; Phelps, D.; Kriz, R.; Recny, M.; Sultzman, L.;
J. Biol. Chem. 261, 9029-9033, 1986
A;Title: Isolation and characterization of cDNA clones for the 35-kDa pulmonary surfacta
A;Reference number: A25720; MUID:86250832; PMID:3755136
A;Accession: A25720
A;Molecule type: mRNA
A;Residues: 1-248 <FLO>
A;Cross-references: GB:M13686; NID:g190669; PIDN:AAA60211.1; PID:g190670
A;Note: part of the sequence was confirmed by protein sequencing
A;Note: the amino end of the mature protein, which was not identified, is partially acet
A;Note: clones corresponding to two different proteins were sequenced. Cotranslational m
C;Genetics:
A;Gene: GDB:SFTPA1; SFTP1; SP-A; SP-A1
A;Cross-references: GDB:119593; OMIM:178630
A;Map position: 10q22-10q23
C;Superfamily: mannose-binding lectin; C-type lectin homology
C;Keywords: acetylated amino end; alveolar proteinosis; calcium; gaseous exchange; glycc
F;1-20/Domain: signal sequence #status predicted <SIG>
F;21-248/Product: pulmonary surfactant protein A #status predicted <MAT>
F;127-246/Domain: C-type lectin homology <LCH>
F;21/Modified site: acetylated amino end (Glu) (in mature form) #status predicted
F;30,33,36,42,54,57,63,76,79,82,91,97/Modified site: 4-hydroxyproline (Pro) #status pred
F;207/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 4.1%; Score 10; DB 1; Length 248;
Best Local Similarity 100.0%; Pred. No. 0.13;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 45 GLPGRDGRDG 54
||||||||||
Db 40 GLPGRDGRDG 49

RESULT 11
LNHUP1
pulmonary surfactant protein A precursor (clone 1A) - human
N;Alternate names: pulmonary surfactant 32K apoprotein; pulmonary surfactant-associated
C;Species: Homo sapiens (man)
C;Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 16-Jul-1999
C;Accession: B25720
R;Floros, J.; Steinbrink, R.; Jacobs, K.; Phelps, D.; Kriz, R.; Recny, M.; Sultzman, L.;
J. Biol. Chem. 261, 9029-9033, 1986
A;Title: Isolation and characterization of cDNA clones for the 35-kDa pulmonary surfacta
A;Reference number: A25720; MUID:86250832; PMID:3755136
A;Accession: B25720
A;Molecule type: mRNA
A;Residues: 1-248 <FLO>
A;Cross-references: GB:K03475
A;Note: part of the sequence was confirmed by protein sequencing
A;Note: the amino end of the mature protein, which was not identified, is partially acet
A;Note: clones corresponding to two different proteins were sequenced. Cotranslational m
C;Genetics:
A;Gene: GDB:SFTPA1; SFTP1; SP-A; SP-A1
A;Cross-references: GDB:119593; OMIM:178630
A;Map position: 10q22-10q23
C;Superfamily: mannose-binding lectin; C-type lectin homology
C;Keywords: acetylated amino end; alveolar proteinosis; calcium; gaseous exchange; glycc
F;1-20/Domain: signal sequence #status predicted <SIG>
F;21-248/Product: pulmonary surfactant protein A #status predicted <MAT>
F;127-246/Domain: C-type lectin homology <LCH>
F;21/Modified site: acetylated amino end (Glu) (in mature form) #status predicted
F;30,33,36,42,54,57,63,76,79,82,91,97/Modified site: 4-hydroxyproline (Pro) #status pred
F;207/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 4.1%; Score 10; DB 1; Length 248;
Best Local Similarity 100.0%; Pred. No. 0.13;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 45 GLPGRDGRDG 54
||||||||||
Db 40 GLPGRDGRDG 49

RESULT 12
LNDGPS
pulmonary surfactant protein A precursor - dog
N;Alternate names: pulmonary surfactant 32K apoprotein; pulmonary surfactant-associat
C;Species: Canis lupus familiaris (dog)
C;Date: 30-Sep-1987 #sequence_revision 30-Sep-1987 #text_change 16-Jul-1999
C;Accession: A25296; A61227; A60142
R;Benson, B.; Hawgood, S.; Schilling, J.; Clements, J.; Damm, D.; Cordell, B.; White,
Proc. Natl. Acad. Sci. U.S.A. 82, 6379-6383, 1985
A;Title: Structure of canine pulmonary surfactant apoprotein: cDNA and complete amino
A;Reference number: A25296; MUID:86016705; PMID:3863100
A;Accession: A25296
A;Molecule type: mRNA
A;Residues: 1-248 <BEN>
A;Note: the authors translated the codon TTG for residue 60 as Pro
A;Note: part of the sequence, including the amino end of the mature protein, was conf
R;Liau, D.F.; Ryan, S.F.
Chem. Phys. Lipids 59, 29-38, 1991
A;Title: Purification of surfactant protein A from dog lung by reconstitution with su
A;Reference number: A61227; MUID:92163993; PMID:1790579
A;Accession: A61227
A;Molecule type: protein
A;Residues: 18-32 <LIA>
R;Ross, G.F.; Meuth, J.; Ohning, B.; Kim, Y.; Whitsett, J.A.
Biochim. Biophys. Acta 870, 267-278, 1986
A;Title: Purification of canine surfactant-associated glycoproteins A. Identification
A;Reference number: A60142; MUID:86159848; PMID:3006781
A;Accession: A60142
A;Molecule type: protein
A;Residues: 24-34;95-101,'X',103-108 <ROS>
R;Patthy, L.
Nature 325, 490, 1987
A;Reference number: A93388; MUID:87115834; PMID:3808053
A;Contents: annotation; animal lectin domain homology
C;Comment: Pulmonary surfactant is a complex of phospholipids and proteins that lower
C;Comment: This protein is a sialoglycoprotein synthesized by alveolar type II cells.
pendent on the presence of calcium ions.
C;Superfamily: mannose-binding lectin; C-type lectin homology
C;Keywords: alveolar proteinosis; calcium; disulfide bond; gaseous exchange; hydroxyp
F;1-17/Domain: signal sequence #status predicted <SIG>
F;18-248/Product: pulmonary surfactant protein A #status experimental <MPT>
F;28-102/Region: collagen-like
F;127-246/Domain: C-type lectin homology <LCH>
F;20,207/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;30/Modified site: 4-hydroxyproline (Pro) #status experimental

Query Match 4.1%; Score 10; DB 1; Length 248;
Best Local Similarity 100.0%; Pred. No. 0.13;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 45 GLPGRDGRDG 54
||||||||||
Db 40 GLPGRDGRDG 49

RESULT 13
LNRTPS
pulmonary surfactant protein A precursor - rat
N;Alternate names: pulmonary surfactant 32K apoprotein; pulmonary surfactant-associat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 16-Jul-1999
C;Accession: A29299; JS0034; S23183
R;Sano, K.; Fisher, J.; Mason, R.J.; Kuroki, Y.; Schilling, J.; Benson, B.; Voelker,
Biochem. Biophys. Res. Commun. 144, 367-374, 1987
A;Title: Isolation and sequence of a cDNA clone for the rat pulmonary surfactant-asso
A;Reference number: A29299; MUID:87213191; PMID:3579914

A;Contents: 0.9 kb cDNA
A;Accession: A29299
A;Molecule type: mRNA
A;Residues: 1-248 <SAN>
A;Note: part of the sequence, including the amino end of the mature protein, was confirm
R;Fisher, J.H.; Emrie, P.A.; Shannon, J.; Sano, K.; Hattler, B.; Mason, R.J.
Biochim. Biophys. Acta 950, 338-345, 1988
A;Title: Rat pulmonary surfactant protein A is expressed as two differently sized mRNA s
A;Reference number: JS0034; MUID:89000785; PMID:2901856
A;Contents: 1.6 kb cDNA
A;Accession: JS0034
A;Molecule type: mRNA
A;Residues: 1-248 <FIS>
A;Note: the codons given for residues 78, 84, and 180 are inconsistent with the authors'
R;Lacaze-Masmonteil, T.; Fraslon, C.; Bourbon, J.; Raymondjean, M.; Kahn, A.
Eur. J. Biochem. 206, 613-623, 1992
A;Title: Characterization of the rat pulmonary surfactant protein A promoter.
A;Reference number: S23183; MUID:92298987; PMID:1606951
A;Accession: S23183
A;Molecule type: DNA
A;Residues: 1-32 <LAC>
C;Comment: Pulmonary surfactant is a complex of phospholipids and proteins that lowers t
C;Comment: This protein is a sialoglycoprotein synthesized by alveolar type II cells. It
pendent on the presence of calcium ions.
C;Comment: Two species of mRNA, which probably are transcribed from a single gene, have
C;Comment: Size heterogeneity of these proteins arises from posttranslational modificati
C;Superfamily: mannose-binding lectin; C-type lectin homology
C;Keywords: alveolar proteinosis; calcium; gaseous exchange; glycoprotein; hydroxyprolin
F;1-20/Domain: signal sequence #status predicted <SIG>
F;21-248/Product: pulmonary surfactant protein A #status experimental <MAT>
F;37-109/Region: collagen-like
F;127-246/Domain: C-type lectin homology <LCH>
F;21/Binding site: carbohydrate (Asn) (covalent) #status absent
F;30,33,36,42,54,57,63,67,70,76/Modified site: 4-hydroxyproline (Pro) #status experiment
F;207/Binding site: carbohydrate (Asn) (covalent) #status predicted

  Query Match 4.1%; Score 10; DB 1; Length 248;
  Best Local Similarity 100.0%; Pred. No. 0.13;
  Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 45 GLPGRDGRDG 54
 ||||||||||
Db 40 GLPGRDGRDG 49

RESULT 14
I51921
pulmonary surfactant-associated protein A1 - human
N;Alternate names: SP-A1
C;Species: Homo sapiens (man)
C;Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 20-Aug-1999
C;Accession: I51921
R;Katyal, S.L.; Singh, G.; Locker, J.
Am. J. Respir. Cell Mol. Biol. 6, 446-452, 1992
A;Title: Characterization of a second human pulmonary surfactant-associated protein SP-A
A;Reference number: I51921; MUID:92198680; PMID:1372511
A;Accession: I51921
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-248 <RES>
A;Cross-references: GB:M68519; NID:g338048; PIDN:AAA60319.1; PID:g338049
C;Genetics:
A;Gene: GDB:SFTPA1; SFTP1; SP-A; SP-A1
A;Cross-references: GDB:119593; OMIM:178630
A;Map position: 10q22-10q23
A;Introns: 58/1; 98/1; 124/1
C;Superfamily: mannose-binding lectin; C-type lectin homology
F;127-246/Domain: C-type lectin homology <LCH>

  Query Match 4.1%; Score 10; DB 2; Length 248;
  Best Local Similarity 100.0%; Pred. No. 0.13;
  Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 45 GLPGRDGRDG 54
 ||||||||||
Db 40 GLPGRDGRDG 49

RESULT 15
A48853
pulmonary surfactant-associated protein SP-A - mouse
C;Species: Mus musculus (house mouse)
C;Date: 19-Dec-1993 #sequence_revision 18-Nov-1994 #text_change 20-Aug-1999
C;Accession: A48853
R;Korfhagen, T.R.; Bruno, M.D.; Glasser, S.W.; Ciraolo, P.J.; Whitsett, J.A.; Lattier
Am. J. Physiol. 263, L546-L554, 1992
A;Title: Murine pulmonary surfactant SP-A gene: cloning, sequence, and transcriptiona
A;Reference number: A48853; MUID:93072386; PMID:1443158
A;Contents: DBA/2J
A;Accession: A48853
A;Status: preliminary
A;Molecule type: nucleic acid
A;Residues: 1-248 <KOR>
A;Cross-references: GB:S48768; NID:g260452; PIDN:AAB24274.1; PID:g260453
A;Note: sequence extracted from NCBI backbone (NCBIN:118740, NCBIP:118741)
C;Superfamily: mannose-binding lectin; C-type lectin homology
F;127-246/Domain: C-type lectin homology <LCH>

  Query Match 4.1%; Score 10; DB 2; Length 248;
  Best Local Similarity 100.0%; Pred. No. 0.13;
  Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 45 GLPGRDGRDG 54
 ||||||||||
Db 40 GLPGRDGRDG 49

Search completed: June 18, 2003, 15:06:31
Job time : 41 secs

OM protein - protein search, using sw model

Run on: June 18, 2003, 14:59:55 ; Search time 81 Seconds
(without alignments)
618.141 Million cell updates/sec

Title: US-09-943-851A-42
Perfect score: 243
Sequence: 1 MRPLLVLLLLGLAAGSPPLD.....DSTFSGFLVYSDWHSSPVFA 243

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 671580 seqs, 206047115 residues

Word size : 0

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : SPTREMBL_21:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 62 | 25.5 | 243 | 11 | Q8R002 | Q8r002 mus musculu |
| 2 | 11 | 4.5 | 27 | 4 | Q9UCJ3 | Q9ucj3 homo sapien |
| 3 | 11 | 4.5 | 289 | 5 | Q19813 | Q19813 caenorhabdi |
| 4 | 11 | 4.5 | 306 | 2 | Q9ACN2 | Q9acn2 streptococc |
| 5 | 11 | 4.5 | 375 | 4 | Q8TCD8 | Q8tcd8 homo sapien |
| 6 | 11 | 4.5 | 378 | 6 | Q9N1X4 | Q9n1x4 sus scrofa |
| 7 | 10 | 4.1 | 57 | 6 | O02842 | O02842 papio cynoc |
| 8 | 10 | 4.1 | 238 | 13 | O57451 | O57451 gallus gall |
| 9 | 10 | 4.1 | 246 | 13 | Q9I9Q7 | Q9i9q7 carassius a |
| 10 | 10 | 4.1 | 248 | 6 | Q9TT06 | Q9tt06 ovis aries |
| 11 | 10 | 4.1 | 248 | 6 | Q9N0G1 | Q9n0g1 equus cabal |
| 12 | 10 | 4.1 | 248 | 6 | Q95L88 | Q95188 equus cabal |
| 13 | 10 | 4.1 | 248 | 11 | Q9CQI1 | Q9cqi1 mus musculu |
| 14 | 10 | 4.1 | 254 | 13 | Q98TA4 | Q98ta4 gallus gall |
| 15 | 10 | 4.1 | 256 | 13 | Q9I9Q9 | Q9i9q9 cyprinus ca |
| 16 | 10 | 4.1 | 282 | 5 | Q8WP36 | Q8wp36 suberites d |
| 17 | 10 | 4.1 | 358 | 12 | Q9J3U4 | Q9j3u4 ectocarpus |
| 18 | 10 | 4.1 | 447 | 12 | Q9YTJ3 | Q9ytj3 ateline her |
| 19 | 10 | 4.1 | 694 | 16 | Q97T45 | Q97t45 streptococc |
| 20 | 10 | 4.1 | 812 | 5 | Q06452 | Q06452 ephydatia m |
| 21 | 10 | 4.1 | 1761 | 5 | O18407 | O18407 drosophila |
| 22 | 10 | 4.1 | 1940 | 5 | Q9VMV5 | Q9vmv5 drosophila |
| 23 | 9 | 3.7 | 182 | 11 | Q9CYS4 | Q9cys4 mus musculu |
| 24 | 9 | 3.7 | 260 | 4 | Q9UHG2 | Q9uhg2 homo sapien |
| 25 | 9 | 3.7 | 260 | 11 | Q9QXU9 | Q9qxu9 rattus norv |
| 26 | 9 | 3.7 | 441 | 13 | Q90YI9 | Q90y19 brachydanio |
| 27 | 9 | 3.7 | 567 | 10 | Q9SWQ3 | Q9swq3 artemisia a |
| 28 | 9 | 3.7 | 589 | 16 | Q9AK64 | Q9ak64 streptomyce |
| 29 | 9 | 3.7 | 684 | 5 | P90679 | P90679 arenicola m |
| 30 | 9 | 3.7 | 695 | 11 | Q64348 | Q64348 mus musculu |
| 31 | 9 | 3.7 | 751 | 11 | Q60709 | Q60709 mus musculu |
| 32 | 9 | 3.7 | 763 | 11 | Q61482 | Q61482 mus musculu |
| 33 | 9 | 3.7 | 1752 | 5 | Q07265 | Q07265 strongyloce |
| 34 | 9 | 3.7 | 2288 | 5 | Q23081 | Q23081 caenorhabdi |
| 35 | 9 | 3.7 | 2315 | 5 | Q95ZK3 | Q95zk3 caenorhabdi |
| 36 | 8 | 3.3 | 46 | 11 | Q63074 | Q63074 rattus norv |
| 37 | 8 | 3.3 | 52 | 16 | Q9HZ09 | Q9hz09 pseudomonas |
| 38 | 8 | 3.3 | 87 | 2 | Q45313 | Q45313 bacillus me |
| 39 | 8 | 3.3 | 107 | 4 | Q9BQY7 | Q9bqy7 homo sapien |
| 40 | 8 | 3.3 | 109 | 11 | Q9CVJ2 | Q9cvj2 mus musculu |
| 41 | 8 | 3.3 | 111 | 16 | Q9RD99 | Q9rd99 streptomyce |
| 42 | 8 | 3.3 | 149 | 16 | Q9JXB1 | Q9jxb1 neisseria m |
| 43 | 8 | 3.3 | 149 | 16 | Q9JWV2 | Q9jwv2 neisseria m |
| 44 | 8 | 3.3 | 164 | 16 | P74713 | P74713 synechocyst |
| 45 | 8 | 3.3 | 170 | 6 | Q9GLV5 | Q9glv5 macaca mula |

ALIGNMENTS

RESULT 1
Q8R002
ID Q8R002 PRELIMINARY; PRT; 243 AA.
AC Q8R002;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Similar to DKFZP586B0621 protein (Hypothetical 25.4 kDa protein).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DDBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.
DR EMBL; BC023068; AAH23068.1; -.
DR EMBL; BC025174; AAH25174.1; -.
KW Hypothetical protein.
SQ SEQUENCE 243 AA; 25420 MW; 498129CD051DB97B CRC64;

Query Match 25.5%; Score 62; DB 11; Length 243;
Best Local Similarity 100.0%; Pred. No. 4.8e-52;
Matches 62; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      182 FGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWHSSPV 241
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      182 FGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWHSSPV 241

Qy      242 FA 243
            ||
Db      242 FA 243
```

RESULT 2

```
Q9UCJ3
ID   Q9UCJ3      PRELIMINARY;      PRT;    27 AA.
AC   Q9UCJ3;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   Surfactant protein D (Fragment).
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=93142849; PubMed=8424457;
RA   Crouch E., Persson A., Chang D.;
RT   "Accumulation of surfactant protein D in human pulmonary alveolar
RT   proteinosis.";
RL   Am. J. Pathol. 142:241-248(1993).
DR   InterPro; IPR000087; Collagen.
SQ   SEQUENCE   27 AA;  2696 MW;  8F770650E2A967CA CRC64;

  Query Match             4.5%;  Score 11;  DB 4;  Length 27;
  Best Local Similarity   100.0%;  Pred. No. 0.0028;
  Matches   11;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         45 GLPGRDGRDGR 55
              ||||||||||||
Db          1 GLPGRDGRDGR 11


RESULT 3
Q19813
ID   Q19813      PRELIMINARY;      PRT;   289 AA.
AC   Q19813;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   F26F12.1 protein.
GN   F26F12.1.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BRISTOL N2;
RX   MEDLINE=99069613; PubMed=9851916;
RA   None;
RT   "Genome sequence of the nematode C. elegans: a platform for
RT   investigating biology. The C. elegans Sequencing Consortium.";
RL   Science 282:2012-2018(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BRISTOL N2;
RA   Wilson R., Bentley D., Gattung S.;
RT   "The sequence of C. elegans cosmid F26F12.";
RL   Submitted (APR-1996) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BRISTOL N2;
RA   Waterston R.;
RL   Submitted (JUN-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U55373; AAC25888.1; -.
DR   InterPro; IPR000087; Collagen.
DR   InterPro; IPR002486; Col_cuticle_N.
DR   Pfam; PF01391; Collagen; 3.
DR   Pfam; PF01484; Col_cuticle_N; 1.
SQ   SEQUENCE   289 AA;  28740 MW;  DFFB78B2B34B29DF CRC64;

  Query Match             4.5%;  Score 11;  DB 5;  Length 289;
  Best Local Similarity   100.0%;  Pred. No. 0.026;
  Matches   11;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         80 PGPRGEAGPAG 90
              ||||||||||||
Db        222 PGPRGEAGPAG 232


RESULT 4
Q9ACN2
ID   Q9ACN2      PRELIMINARY;      PRT;   306 AA.
AC   Q9ACN2;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   SclB protein.
GN   SCLB.
OS   Streptococcus pyogenes.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;
OC   Streptococcaceae; Streptococcus.
OX   NCBI_TaxID=1314;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=655;
RX   MEDLINE=21097281; PubMed=11158359;
RA   Whatmore A.M.;
RT   "Streptococcus pyogenes sclB encodes a putative hypervariable surface
RT   protein with a collagen-like repetitive structure.";
RL   Microbiology 147:419-429(2001).
DR   EMBL; AJ301809; CAC33778.1; -.
DR   InterPro; IPR000087; Collagen.
DR   InterPro; IPR001899; Gram_pos_anchor.
DR   Pfam; PF01391; Collagen; 1.
DR   Pfam; PF00746; Gram_pos_anchor; 1.
DR   PRINTS; PR00015; GPOSANCHOR.
DR   TIGRFAMs; TIGR01167; LPXTG_anchor; 1.
DR   PROSITE; PS00343; GRAM_POS_ANCHORING; UNKNOWN_1.
KW   Transmembrane.
SQ   SEQUENCE   306 AA;  31833 MW;  57EBDC244DF17600 CRC64;

  Query Match             4.5%;  Score 11;  DB 2;  Length 306;
  Best Local Similarity   100.0%;  Pred. No. 0.027;
  Matches   11;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         81 GPRGEAGPAGP 91
              ||||||||||||
Db        148 GPRGEAGPAGP 158


RESULT 5
Q8TCD8
ID   Q8TCD8      PRELIMINARY;      PRT;   375 AA.
AC   Q8TCD8;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Hypothetical 37.7 kDa protein.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=LUNG;
RA   Strausberg R.;
RL   Submitted (FEB-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC022318; AAH22318.1; -.
KW   Hypothetical protein.
SQ   SEQUENCE   375 AA;  37655 MW;  CCB7375D3C86421A CRC64;

  Query Match             4.5%;  Score 11;  DB 4;  Length 375;
  Best Local Similarity   100.0%;  Pred. No. 0.033;
  Matches   11;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         45 GLPGRDGRDGR 55
              ||||||||||||
```

```
Db          46 GLPGRDGRDGR 56


RESULT 6
Q9N1X4
ID   Q9N1X4       PRELIMINARY;       PRT;   378 AA.
AC   Q9N1X4;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Lung surfactant protein D precursor.
GN   SFTPD.
OS   Sus scrofa (Pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX   NCBI_TaxID=9823;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20109098; PubMed=10640760;
RA   van Eijk M., Haagsman H.P., Skinner T., Archibald A.L., Reid K.B.M.,
RA   Lawson P.R.;
RT   "Porcine Lung Surfactant Protein D (SP-D): cDNA cloning, chromosomal
RT   localisation and tissue distribution.";
RL   J. Immunol. 164:1442-1450(2000).
DR   EMBL; AF132496; AAF22145.2; -.
DR   HSSP; P35247; 1B08.
DR   InterPro; IPR000087; Collagen.
DR   InterPro; IPR001304; Lectin_C.
DR   Pfam; PF01391; Collagen; 3.
DR   Pfam; PF00059; lectin_c; 1.
DR   ProDom; PD000007; Collagen; 2.
DR   SMART; SM00034; CLECT; 1.
DR   PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
KW   Signal.
FT   SIGNAL        1     20       POTENTIAL.
FT   CHAIN        21    378       LUNG SURFACTANT PROTEIN D.
SQ   SEQUENCE   378 AA;  37986 MW;  3504E8C1E56C341D CRC64;

  Query Match             4.5%;  Score 11;  DB 6;  Length 378;
  Best Local Similarity   100.0%;  Pred. No. 0.033;
  Matches   11;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          45 GLPGRDGRDGR 55
               ||||||||||||
Db          46 GLPGRDGRDGR 56


RESULT 7
O02842
ID   O02842       PRELIMINARY;       PRT;    57 AA.
AC   O02842;
DT   01-JUL-1997 (TrEMBLrel. 04, Created)
DT   01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Surfactant protein A2 (Fragment).
OS   Papio cynocephalus (Yellow baboon).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC   Cercopithecinae; Papio.
OX   NCBI_TaxID=9556;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=KIDNEY;
RX   MEDLINE=97053398; PubMed=8897910;
RA   Gao E., Wang Y., McCormick S.M., Li J., Seidner S.R., Mendelson C.R.;
RT   "Characterization of two baboon surfactant protein A genes.";
RL   Am. J. Physiol. 271:L617-L630(1996).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=KIDNEY;
RA   Gao E., Wang Y., McCormick S.M., Li J., Seidner S.R., Mendelson C.R.;
RL   Submitted (MAY-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF003897; AAB61295.1; -.
DR   EMBL; AF003896; AAB61294.1; -.
DR   InterPro; IPR000087; Collagen.
DR   ProDom; PD000007; Collagen; 1.
FT   NON_TER      57     57
SQ   SEQUENCE   57 AA;  5584 MW;  B25149EC822F643C CRC64;

  Query Match             4.1%;  Score 10;  DB 6;  Length 57;
  Best Local Similarity   100.0%;  Pred. No. 0.053;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          45 GLPGRDGRDG 54
               ||||||||||
Db          40 GLPGRDGRDG 49


RESULT 8
O57451
ID   O57451       PRELIMINARY;       PRT;   238 AA.
AC   O57451;
DT   01-JUN-1998 (TrEMBLrel. 06, Created)
DT   01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Mannan-binding lectin (Fragment).
GN   CMBL.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=WHITE LEGHORN; TISSUE=LIVER;
RA   Laursen S.B.;
RL   Submitted (AUG-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF022226; AAB94071.1; -.
DR   HSSP; P19999; 1YTT.
DR   InterPro; IPR000087; Collagen.
DR   InterPro; IPR001304; Lectin_C.
DR   Pfam; PF01391; Collagen; 1.
DR   Pfam; PF00059; lectin_c; 1.
DR   ProDom; PD000007; Collagen; 1.
DR   SMART; SM00034; CLECT; 1.
DR   PROSITE; PS00615; C_TYPE_LECTIN_1; UNKNOWN_1.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
KW   Lectin.
FT   NON_TER     238    238
SQ   SEQUENCE   238 AA;  25645 MW;  E5C9B5197AAE64E3 CRC64;

  Query Match             4.1%;  Score 10;  DB 13;  Length 238;
  Best Local Similarity   100.0%;  Pred. No. 0.2;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          45 GLPGRDGRDG 54
               ||||||||||
Db          31 GLPGRDGRDG 40


RESULT 9
Q9I9Q7
ID   Q9I9Q7       PRELIMINARY;       PRT;   246 AA.
AC   Q9I9Q7;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Mannose binding-like lectin precursor (Fragment).
GN   MBL.
OS   Carassius auratus (Goldfish).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC   Cyprinidae; Carassius.
OX   NCBI_TaxID=7957;
RN   [1]
```

```
RP   SEQUENCE FROM N.A.
RC   TISSUE=LIVER;
RX   MEDLINE=20456722; PubMed=11003389;
RA   Vitved L., Holmskov U., Koch C., Teisner B., Hansen S., Skjodt K.;
RT   "The homologue of mannose-binding lectin in the carp family Cyprinidae
RT   is expressed at high level in spleen, and the deduced primary
RT   structure predicts affinity for galactose.";
RL   Immunogenetics 51:955-964(2000).
DR   EMBL; AF227739; AAF63470.1; -.
DR   HSSP; P35247; 1B08.
DR   InterPro; IPR000087; Collagen.
DR   InterPro; IPR001304; Lectin_C.
DR   Pfam; PF01391; Collagen; 2.
DR   Pfam; PF00059; lectin_c; 1.
DR   SMART; SM00034; CLECT; 1.
DR   PROSITE; PS00615; C_TYPE_LECTIN_1; UNKNOWN_1.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
KW   Lectin; Signal.
FT   NON_TER       1      1
FT   SIGNAL        <1     13       POTENTIAL.
FT   VARIANT       145    145      S -> F.
SQ   SEQUENCE   246 AA;  25709 MW;  AB692282D289D0D5 CRC64;

  Query Match             4.1%;  Score 10;  DB 13;  Length 246;
  Best Local Similarity   100.0%;  Pred. No. 0.21;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         45 GLPGRDGRDG 54
              ||||||||||
Db         35 GLPGRDGRDG 44


RESULT 10
Q9TT06
ID   Q9TT06      PRELIMINARY;      PRT;   248 AA.
AC   Q9TT06;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Pulmonary surfactant protein A (Pulmonary surfactant-associated
DE   protein A).
GN   SPAS OR SP-A.
OS   Ovis aries (Sheep).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Caprinae; Ovis.
OX   NCBI_TaxID=9940;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20215262; PubMed=10749753;
RA   Braems G.A., Yao L.-J., Inchley K., Brickenden A., Han V.K.M.,
RA   Grolla A., Challis J.R.G., Possmayer F.;
RT   "Ovine surfactant protein cDNAs: use in studies on fetal lung growth
RT   and maturation after prolonged hypoxemia.";
RL   Am. J. Physiol. 278:L754-L764(2000).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=LUNG;
RX   MEDLINE=20215263; PubMed=10749754;
RA   Pietschmann S.M., Pison U.;
RT   "cDNA cloning of ovine pulmonary SP-A, SP-B, and SP-C: isolation of
RT   two different sequences for SP-B.";
RL   Am. J. Physiol. 278:L765-L778(2000).
DR   EMBL; AF211856; AAF18995.1; -.
DR   EMBL; AF076633; AAF31148.1; -.
DR   HSSP; P22897; 1EGG.
DR   InterPro; IPR000087; Collagen.
DR   InterPro; IPR001304; Lectin_C.
DR   Pfam; PF01391; Collagen; 2.
DR   Pfam; PF00059; lectin_c; 1.
DR   SMART; SM00034; CLECT; 1.
DR   PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
SQ   SEQUENCE   248 AA;  26394 MW;  D65E7293BBFF1FD9 CRC64;

  Query Match             4.1%;  Score 10;  DB 6;  Length 248;
  Best Local Similarity   100.0%;  Pred. No. 0.21;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         45 GLPGRDGRDG 54
              ||||||||||
Db         40 GLPGRDGRDG 49


RESULT 11
Q9N0G1
ID   Q9N0G1      PRELIMINARY;      PRT;   248 AA.
AC   Q9N0G1;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Pulmonary surfactant protein A.
OS   Equus caballus (Horse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
OX   NCBI_TaxID=9796;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=THOROUGHBRED; TISSUE=LUNG;
RA   Hobo S.;
RT   "Molecular cloning of equine pulmonary surfactant proteins.";
RL   Submitted (JUL-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB015963; BAA97976.1; -.
DR   HSSP; P35247; 1B08.
DR   InterPro; IPR000087; Collagen.
DR   InterPro; IPR001304; Lectin_C.
DR   Pfam; PF01391; Collagen; 2.
DR   Pfam; PF00059; lectin_c; 1.
DR   SMART; SM00034; CLECT; 1.
DR   PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
SQ   SEQUENCE   248 AA;  26000 MW;  BBE12EFB05C2B8D1 CRC64;

  Query Match             4.1%;  Score 10;  DB 6;  Length 248;
  Best Local Similarity   100.0%;  Pred. No. 0.21;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         45 GLPGRDGRDG 54
              ||||||||||
Db         40 GLPGRDGRDG 49


RESULT 12
Q95L88
ID   Q95L88      PRELIMINARY;      PRT;   248 AA.
AC   Q95L88;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Pulmonary surfactant-associated protein A.
OS   Equus caballus (Horse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
OX   NCBI_TaxID=9796;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Weber B.I.L., Hospes R., Gortner L.;
RL   Submitted (JUL-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF400580; AAL07690.1; -.
DR   InterPro; IPR000087; Collagen.
DR   InterPro; IPR001304; Lectin_C.
DR   Pfam; PF01391; Collagen; 2.
DR   Pfam; PF00059; lectin_c; 1.
DR   PROSITE; PS00615; C_TYPE_LECTIN_1; UNKNOWN_1.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
SQ   SEQUENCE   248 AA;  26047 MW;  B71133E005C9A5C1 CRC64;
```

```
  Query Match             4.1%;  Score 10;  DB 6;  Length 248;
  Best Local Similarity   100.0%;  Pred. No. 0.21;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         45 GLPGRDGRDG 54
              ||||||||||
Db         40 GLPGRDGRDG 49


RESULT 13
Q9CQI1
ID   Q9CQI1      PRELIMINARY;      PRT;   248 AA.
AC   Q9CQI1;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Surfactant associated protein A.
GN   SFTPA.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=EMBRYO, AND LUNG;
RX   MEDLINE=21085660; PubMed=11217851;
RA   Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA   Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA   Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA   Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA   Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA   Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA   Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA   Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA   Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA   Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA   Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA   Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA   Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA   Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA   Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA   Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA   Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA   Hayashizaki Y.;
RT   "Functional annotation of a full-length mouse cDNA collection.";
RL   Nature 409:685-690(2001).
DR   EMBL; AK011333; BAB27551.1; -.
DR   EMBL; AK004620; BAB23416.1; -.
DR   EMBL; AK004788; BAB23565.1; -.
DR   HSSP; P35247; 1B08.
DR   MGD; MGI:109518; Sftpa.
DR   InterPro; IPR000087; Collagen.
DR   InterPro; IPR001304; Lectin_C.
DR   Pfam; PF01391; Collagen; 2.
DR   Pfam; PF00059; lectin_c; 1.
DR   SMART; SM00034; CLECT; 1.
DR   PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
SQ   SEQUENCE   248 AA;  26183 MW;  8A5670CFAD3EB9B6 CRC64;

  Query Match             4.1%;  Score 10;  DB 11;  Length 248;
  Best Local Similarity   100.0%;  Pred. No. 0.21;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         45 GLPGRDGRDG 54
              ||||||||||
Db         40 GLPGRDGRDG 49


RESULT 14
Q98TA4
ID   Q98TA4      PRELIMINARY;      PRT;   254 AA.
AC   Q98TA4;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Mannose-binding lectin protein precursor.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20456722; PubMed=11003389;
RA   Vitved L., Holmskov U., Koch C., Teisner B., Hansen S., Skjodt K.;
RT   "The homologue of mannose-binding lectin in the carp family Cyprinidae
RT   is expressed at high level in spleen, and the deduced primary
RT   structure predicts affinity for galactose.";
RL   Immunogenetics 51:955-964(2000).
DR   EMBL; AF231714; AAK30298.1; -.
DR   HSSP; P19999; 1YTT.
DR   InterPro; IPR000087; Collagen.
DR   InterPro; IPR000005; HTHAraC.
DR   InterPro; IPR001304; Lectin_C.
DR   Pfam; PF01391; Collagen; 1.
DR   Pfam; PF00059; lectin_c; 1.
DR   ProDom; PD000007; Collagen; 1.
DR   SMART; SM00034; CLECT; 1.
DR   PROSITE; PS00615; C_TYPE_LECTIN_1; UNKNOWN_1.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
DR   PROSITE; PS00041; HTH_ARAC_FAMILY_1; UNKNOWN_1.
KW   Lectin; Signal.
FT   SIGNAL        1     21       POTENTIAL.
FT   CHAIN        22    254       MANNOSE-BINDING LECTIN PROTEIN.
FT   VARIANT     234    234       D -> V.
SQ   SEQUENCE   254 AA;  27376 MW;  C924428643441AED CRC64;

  Query Match             4.1%;  Score 10;  DB 13;  Length 254;
  Best Local Similarity   100.0%;  Pred. No. 0.21;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         45 GLPGRDGRDG 54
              ||||||||||
Db         47 GLPGRDGRDG 56


RESULT 15
Q9I9Q9
ID   Q9I9Q9      PRELIMINARY;      PRT;   256 AA.
AC   Q9I9Q9;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Mannose binding-like lectin precursor.
GN   MBL.
OS   Cyprinus carpio (Common carp).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC   Cyprinidae; Cyprinus.
OX   NCBI_TaxID=7962;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=LIVER;
RX   MEDLINE=20456722; PubMed=11003389;
RA   Vitved L., Holmskov U., Koch C., Teisner B., Hansen S., Skjodt K.;
RT   "The homologue of mannose-binding lectin in the carp family Cyprinidae
RT   is expressed at high level in spleen, and the deduced primary
RT   structure predicts affinity for galactose.";
RL   Immunogenetics 51:955-964(2000).
DR   EMBL; AF227737; AAF63468.1; -.
DR   HSSP; P35247; 1B08.
DR   InterPro; IPR000087; Collagen.
DR   InterPro; IPR001304; Lectin_C.
DR   Pfam; PF01391; Collagen; 2.
```

```
DR   Pfam; PF00059; lectin_c; 1.
DR   SMART; SM00034; CLECT; 1.
DR   PROSITE; PS00615; C_TYPE_LECTIN_1; UNKNOWN_1.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
KW   Lectin; Signal.
FT   SIGNAL        1     23       POTENTIAL.
FT   CHAIN        24    256       MANNOSE BINDING-LIKE LECTIN.
FT   VARIANT     235    235       S -> T.
SQ   SEQUENCE   256 AA;  26934 MW;  D019291D1167730D CRC64;

  Query Match             4.1%;  Score 10;  DB 13;  Length 256;
  Best Local Similarity   100.0%;  Pred. No. 0.22;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         45 GLPGRDGRDG 54
              ||||||||||
Db         45 GLPGRDGRDG 54


Search completed: June 18, 2003, 15:05:52
Job time : 91 secs
```

GenCore version 5.1.6


OM protein - protein search, using sw model

Run on: June 18, 2003, 15:02:44 ; Search time 26 Seconds
(without alignments)
274.991 Million cell updates/sec

Title: US-09-943-851A-42
Perfect score: 243
Sequence: 1 MRPLLVLLLLGLAAGSPPLD.....DSTFSGFLVYSDWHSSPVFA 243

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 262574 seqs, 29422922 residues

Word size : 0

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 243 | 100.0 | 243 | 4 | US-09-140-804-2 | Sequence 2, Appli |
| 2 | 243 | 100.0 | 243 | 4 | US-09-336-536-3 | Sequence 3, Appli |
| 3 | 228 | 93.8 | 228 | 4 | US-09-336-536-4 | Sequence 4, Appli |
| 4 | 128 | 52.7 | 128 | 4 | US-09-336-536-7 | Sequence 7, Appli |
| 5 | 70 | 28.8 | 243 | 4 | US-09-188-930-295 | Sequence 295, App |
| 6 | 62 | 25.5 | 228 | 4 | US-09-336-536-11 | Sequence 11, Appl |
| 7 | 62 | 25.5 | 243 | 4 | US-09-336-536-10 | Sequence 10, Appl |
| 8 | 60 | 24.7 | 60 | 4 | US-09-336-536-6 | Sequence 6, Appli |
| 9 | 51 | 21.0 | 128 | 4 | US-09-336-536-14 | Sequence 14, Appl |
| 10 | 43 | 17.7 | 60 | 4 | US-09-336-536-13 | Sequence 13, Appl |
| 11 | 15 | 6.2 | 15 | 4 | US-09-336-536-5 | Sequence 5, Appli |
| 12 | 11 | 4.5 | 128 | 4 | US-09-227-357-190 | Sequence 190, App |
| 13 | 10 | 4.1 | 1057 | 3 | US-08-931-820-4 | Sequence 4, Appli |
| 14 | 10 | 4.1 | 1078 | 3 | US-08-963-825-21 | Sequence 21, Appl |
| 15 | 10 | 4.1 | 1078 | 4 | US-09-500-811-21 | Sequence 21, Appl |
| 16 | 10 | 4.1 | 1078 | 4 | US-09-570-573-21 | Sequence 21, Appl |
| 17 | 10 | 4.1 | 1078 | 4 | US-09-548-608-21 | Sequence 21, Appl |
| 18 | 8 | 3.3 | 130 | 4 | US-09-485-316A-10 | Sequence 10, Appl |
| 19 | 8 | 3.3 | 170 | 1 | US-08-313-681A-2 | Sequence 2, Appli |
| 20 | 8 | 3.3 | 170 | 3 | US-09-322-911-2 | Sequence 2, Appli |
| 21 | 8 | 3.3 | 184 | 1 | US-08-211-942-7 | Sequence 7, Appli |
| 22 | 8 | 3.3 | 184 | 1 | US-08-211-942-15 | Sequence 15, Appl |
| 23 | 8 | 3.3 | 186 | 1 | US-08-211-942-9 | Sequence 9, Appli |
| 24 | 8 | 3.3 | 222 | 4 | US-09-140-804-7 | Sequence 7, Appli |
| 25 | 8 | 3.3 | 231 | 4 | US-09-530-423-2 | Sequence 2, Appli |
| 26 | 8 | 3.3 | 244 | 2 | US-08-463-911-7 | Sequence 7, Appli |
| 27 | 8 | 3.3 | 244 | 4 | US-09-140-804-3 | Sequence 3, Appli |
| 28 | 8 | 3.3 | 244 | 4 | US-09-336-536-20 | Sequence 20, Appl |
| 29 | 8 | 3.3 | 244 | 4 | US-09-530-423-1 | Sequence 1, Appli |
| 30 | 8 | 3.3 | 247 | 2 | US-08-463-911-2 | Sequence 2, Appli |
| 31 | 8 | 3.3 | 247 | 4 | US-09-140-804-8 | Sequence 8, Appli |
| 32 | 8 | 3.3 | 247 | 4 | US-09-118-408-3 | Sequence 3, Appli |
| 33 | 8 | 3.3 | 247 | 4 | US-09-506-855-3 | Sequence 3, Appli |
| 34 | 8 | 3.3 | 368 | 1 | US-08-211-942-17 | Sequence 17, Appl |
| 35 | 8 | 3.3 | 547 | 1 | US-08-494-168-7 | Sequence 7, Appli |
| 36 | 8 | 3.3 | 1024 | 3 | US-08-931-820-2 | Sequence 2, Appli |
| 37 | 8 | 3.3 | 1366 | 3 | US-08-963-825-19 | Sequence 19, Appl |
| 38 | 8 | 3.3 | 1366 | 4 | US-09-500-811-19 | Sequence 19, Appl |
| 39 | 8 | 3.3 | 1366 | 4 | US-09-570-573-19 | Sequence 19, Appl |
| 40 | 8 | 3.3 | 1366 | 4 | US-09-548-608-19 | Sequence 19, Appl |
| 41 | 8 | 3.3 | 1366 | 4 | US-09-585-887-10 | Sequence 10, Appl |
| 42 | 8 | 3.3 | 1366 | 4 | US-09-289-578-10 | Sequence 10, Appl |
| 43 | 8 | 3.3 | 1442 | 2 | US-08-316-650-12 | Sequence 12, Appl |
| 44 | 8 | 3.3 | 1442 | 5 | PCT-US95-02251-12 | Sequence 12, Appl |
| 45 | 8 | 3.3 | 1694 | 1 | US-08-494-168-2 | Sequence 2, Appli |

ALIGNMENTS

RESULT 1
US-09-140-804-2
; Sequence 2, Application US/09140804
; Patent No. 6197930
; GENERAL INFORMATION:
; APPLICANT: Sheppard, Paul O.
; APPLICANT: Humes, Jacqueline M.
; TITLE OF INVENTION: ADIPOCYTE-SPECIFIC PROTEIN HOMOLOGS
; FILE REFERENCE: 97-49
; CURRENT APPLICATION NUMBER: US/09/140,804
; CURRENT FILING DATE: 1998-08-26
; EARLIER APPLICATION NUMBER: 60/056,983
; EARLIER FILING DATE: 1997-08-26
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 243
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-140-804-2

Query Match 100.0%; Score 243; DB 4; Length 243;
Best Local Similarity 100.0%; Pred. No. 1.6e-218;
Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MRPLLVLLLLGLAAGSPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MRPLLVLLLLGLAAGSPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPG 60

Qy         61 APGEKGEGGRPGLPGPRGDPGPRGEAGPAGPTGPAGECSVPPRSAFSAKRSESRVPPPSD 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 APGEKGEGGRPGLPGPRGDPGPRGEAGPAGPTGPAGECSVPPRSAFSAKRSESRVPPPSD 120

Qy        121 APLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGESIASFFQ 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 APLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGESIASFFQ 180

Qy        181 FFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWHSSP 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 FFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWHSSP 240

Qy        241 VFA 243
              |||
Db        241 VFA 243
```

RESULT 2
US-09-336-536-3
; Sequence 3, Application US/09336536
; Patent No. 6406884

```
; GENERAL INFORMATION:
;  APPLICANT: Leiby, K.
;  APPLICANT: McKay, C.
;  APPLICANT: Bossone, S.
;  TITLE OF INVENTION: SECRETED PROTEINS AND USES THEREOF
;  FILE REFERENCE: 7853-144
;  CURRENT APPLICATION NUMBER: US/09/336,536
;  CURRENT FILING DATE: 1999-06-18
;  NUMBER OF SEQ ID NOS: 75
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
;   LENGTH: 243
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-336-536-3

  Query Match           100.0%;  Score 243;  DB 4;  Length 243;
  Best Local Similarity  100.0%;  Pred. No. 1.6e-218;
  Matches  243;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MRPLLVLLLLGLAAGSPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MRPLLVLLLLGLAAGSPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPG 60

Qy         61 APGEKGEGGRPGLPGPRGDPGPRGEAGPAGPTGPAGECSVPPRSAFSAKRSESRVPPPSD 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 APGEKGEGGRPGLPGPRGDPGPRGEAGPAGPTGPAGECSVPPRSAFSAKRSESRVPPPSD 120

Qy        121 APLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGESIASFFQ 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 APLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGESIASFFQ 180

Qy        181 FFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWHSSP 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 FFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWHSSP 240

Qy        241 VFA 243
              |||
Db        241 VFA 243


RESULT 3
US-09-336-536-4
; Sequence 4, Application US/09336536
; Patent No. 6406884
; GENERAL INFORMATION:
;  APPLICANT: Leiby, K.
;  APPLICANT: McKay, C.
;  APPLICANT: Bossone, S.
;  TITLE OF INVENTION: SECRETED PROTEINS AND USES THEREOF
;  FILE REFERENCE: 7853-144
;  CURRENT APPLICATION NUMBER: US/09/336,536
;  CURRENT FILING DATE: 1999-06-18
;  NUMBER OF SEQ ID NOS: 75
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
;   LENGTH: 228
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-336-536-4

  Query Match           93.8%;  Score 228;  DB 4;  Length 228;
  Best Local Similarity  100.0%;  Pred. No. 1.4e-204;
  Matches  228;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         16 SPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPGAPGEKGEGGRPGLPG 75
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 SPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPGAPGEKGEGGRPGLPG 60

Qy         76 PRGDPGPRGEAGPAGPTGPAGECSVPPRSAFSAKRSESRVPPPSDAPLPFDRVLVNEQGH 135
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 PRGDPGPRGEAGPAGPTGPAGECSVPPRSAFSAKRSESRVPPPSDAPLPFDRVLVNEQGH 120

Qy        136 YDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGESIASFFQFFGGWPKPASLSGGA 195
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 YDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGESIASFFQFFGGWPKPASLSGGA 180

Qy        196 MVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWHSSPVFA 243
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 MVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWHSSPVFA 228


RESULT 4
US-09-336-536-7 .
; Sequence 7, Application US/09336536
; Patent No. 6406884
; GENERAL INFORMATION:
;  APPLICANT: Leiby, K.
;  APPLICANT: McKay, C.
;  APPLICANT: Bossone, S.
;  TITLE OF INVENTION: SECRETED PROTEINS AND USES THEREOF
;  FILE REFERENCE: 7853-144
;  CURRENT APPLICATION NUMBER: US/09/336,536
;  CURRENT FILING DATE: 1999-06-18
;  NUMBER OF SEQ ID NOS: 75
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
;   LENGTH: 128
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-336-536-7

  Query Match           52.7%;  Score 128;  DB 4;  Length 128;
  Best Local Similarity  100.0%;  Pred. No. 9.7e-112;
  Matches  128;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        105 AFSAKRSESRVPPPSDAPLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASL 164
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AFSAKRSESRVPPPSDAPLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASL 60

Qy        165 QFDLVKNGESIASFFQFFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTD 224
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 QFDLVKNGESIASFFQFFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTD 120

Qy        225 STFSGFLV 232
              ||||||||
Db        121 STFSGFLV 128


RESULT 5
US-09-188-930-295
; Sequence 295, Application US/09188930A
; Patent No. 6150502
; GENERAL INFORMATION:
;  APPLICANT: Watson, James D.
;  APPLICANT: Strachan, Lorna
;  APPLICANT: Sleeman, Matthew
;  APPLICANT: Onrust, Rene
;  APPLICANT: Murison, James Greg
;  TITLE OF INVENTION: Compositions Isolated From Skin Cells
;  TITLE OF INVENTION: and Methods For Their Use
;  FILE REFERENCE: 11000.1011c1
;  CURRENT APPLICATION NUMBER: US/09/188,930A
;  CURRENT FILING DATE: 1998-11-09
;  NUMBER OF SEQ ID NOS: 348
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 295
;   LENGTH: 243
;   TYPE: PRT
;   ORGANISM: Rat
US-09-188-930-295

  Query Match           28.8%;  Score 70;  DB 4;  Length 243;
  Best Local Similarity  100.0%;  Pred. No. 1.6e-57;
  Matches   70;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy        174 SIASFFQFFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVY 233
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        174 SIASFFQFFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVY 233

Qy        234 SDWHSSPVFA 243
              ||||||||||
Db        234 SDWHSSPVFA 243
```

RESULT 6
US-09-336-536-11
; Sequence 11, Application US/09336536
; Patent No. 6406884
; GENERAL INFORMATION:
; APPLICANT: Leiby, K.
; APPLICANT: McKay, C.
; APPLICANT: Bossone, S.
; TITLE OF INVENTION: SECRETED PROTEINS AND USES THEREOF
; FILE REFERENCE: 7853-144
; CURRENT APPLICATION NUMBER: US/09/336,536
; CURRENT FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
; LENGTH: 228
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-336-536-11

```
  Query Match           25.5%;  Score 62;  DB 4;  Length 228;
  Best Local Similarity  100.0%;  Pred. No. 4.2e-50;
  Matches   62;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        182 FGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWHSSPV 241
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        167 FGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWHSSPV 226

Qy        242 FA 243
              ||
Db        227 FA 228
```

RESULT 7
US-09-336-536-10
; Sequence 10, Application US/09336536
; Patent No. 6406884
; GENERAL INFORMATION:
; APPLICANT: Leiby, K.
; APPLICANT: McKay, C.
; APPLICANT: Bossone, S.
; TITLE OF INVENTION: SECRETED PROTEINS AND USES THEREOF
; FILE REFERENCE: 7853-144
; CURRENT APPLICATION NUMBER: US/09/336,536
; CURRENT FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10
; LENGTH: 243
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-336-536-10

```
  Query Match           25.5%;  Score 62;  DB 4;  Length 243;
  Best Local Similarity  100.0%;  Pred. No. 4.5e-50;
  Matches   62;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        182 FGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWHSSPV 241
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        182 FGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWHSSPV 241

Qy        242 FA 243
              ||
Db        242 FA 243
```

RESULT 8
US-09-336-536-6
; Sequence 6, Application US/09336536
; Patent No. 6406884
; GENERAL INFORMATION:
; APPLICANT: Leiby, K.
; APPLICANT: McKay, C.
; APPLICANT: Bossone, S.
; TITLE OF INVENTION: SECRETED PROTEINS AND USES THEREOF
; FILE REFERENCE: 7853-144
; CURRENT APPLICATION NUMBER: US/09/336,536
; CURRENT FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 60
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-336-536-6

```
  Query Match           24.7%;  Score 60;  DB 4;  Length 60;
  Best Local Similarity  100.0%;  Pred. No. 9.6e-49;
  Matches   60;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         36 GTPGHHGSQGLPGRDGRDGRDGAPGAPGEKGEGGRPGLPGPRGDPGPRGEAGPAGPTGPA 95
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GTPGHHGSQGLPGRDGRDGRDGAPGAPGEKGEGGRPGLPGPRGDPGPRGEAGPAGPTGPA 60
```

RESULT 9
US-09-336-536-14
; Sequence 14, Application US/09336536
; Patent No. 6406884
; GENERAL INFORMATION:
; APPLICANT: Leiby, K.
; APPLICANT: McKay, C.
; APPLICANT: Bossone, S.
; TITLE OF INVENTION: SECRETED PROTEINS AND USES THEREOF
; FILE REFERENCE: 7853-144
; CURRENT APPLICATION NUMBER: US/09/336,536
; CURRENT FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14
; LENGTH: 128
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-336-536-14

```
  Query Match           21.0%;  Score 51;  DB 4;  Length 128;
  Best Local Similarity  100.0%;  Pred. No. 4.4e-40;
  Matches   51;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        182 FGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLV 232
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         78 FGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLV 128
```

RESULT 10
US-09-336-536-13
; Sequence 13, Application US/09336536
; Patent No. 6406884
; GENERAL INFORMATION:
; APPLICANT: Leiby, K.
; APPLICANT: McKay, C.
; APPLICANT: Bossone, S.
; TITLE OF INVENTION: SECRETED PROTEINS AND USES THEREOF
; FILE REFERENCE: 7853-144
; CURRENT APPLICATION NUMBER: US/09/336,536
; CURRENT FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 75

```
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 13
;   LENGTH: 60
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-09-336-536-13

  Query Match             17.7%;  Score 43;  DB 4;  Length 60;
  Best Local Similarity   100.0%;  Pred. No. 6.3e-33;
  Matches   43;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         36 GTPGHHGSQGLPGRDGRDGRDGAPGAPGEKGEGGRPGLPGPRG 78
              |||||||||||||||||||||||||||||||||||||||||||
Db          1 GTPGHHGSQGLPGRDGRDGRDGAPGAPGEKGEGGRPGLPGPRG 43


RESULT 11
US-09-336-536-5
; Sequence 5, Application US/09336536
; Patent No. 6406884
; GENERAL INFORMATION:
;  APPLICANT: Leiby, K.
;  APPLICANT: McKay, C.
;  APPLICANT: Bossone, S.
;  TITLE OF INVENTION: SECRETED PROTEINS AND USES THEREOF
;  FILE REFERENCE: 7853-144
;  CURRENT APPLICATION NUMBER: US/09/336,536
;  CURRENT FILING DATE: 1999-06-18
;  NUMBER OF SEQ ID NOS: 75
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
;   LENGTH: 15
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-336-536-5

  Query Match             6.2%;  Score 15;  DB 4;  Length 15;
  Best Local Similarity   100.0%;  Pred. No. 2.1e-07;
  Matches   15;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MRPLLVLLLLGLAAG 15
              |||||||||||||||
Db          1 MRPLLVLLLLGLAAG 15


RESULT 12
US-09-227-357-190
; Sequence 190, Application US/09227357
; Patent No. 6342581
; GENERAL INFORMATION:
;  APPLICANT: Fischer et al.
;  TITLE OF INVENTION: 123 Human Secreted Proteins
;  FILE REFERENCE: PZ010P1
;  CURRENT APPLICATION NUMBER: US/09/227,357
;  CURRENT FILING DATE: 1999-01-08
;  EARLIER APPLICATION NUMBER: PCT/US98/13684
;  EARLIER FILING DATE: 1998-07-07
;  EARLIER APPLICATION NUMBER: 60/051,926
;  EARLIER FILING DATE: 1997-07-08
;  EARLIER APPLICATION NUMBER: 60/052,793
;  EARLIER FILING DATE: 1997-07-08
;  EARLIER APPLICATION NUMBER: 60/051,925
;  EARLIER FILING DATE: 1997-07-08
;  EARLIER APPLICATION NUMBER: 60/051,929
;  EARLIER FILING DATE: 1997-07-08
;  EARLIER APPLICATION NUMBER: 60/052,803
;  EARLIER FILING DATE: 1997-07-08
;  EARLIER APPLICATION NUMBER: 60/052,732
;  EARLIER FILING DATE: 1997-07-08
;  EARLIER APPLICATION NUMBER: 60/051,931
;  EARLIER FILING DATE: 1997-07-08
;  EARLIER APPLICATION NUMBER: 60/051,932
;  EARLIER FILING DATE: 1997-07-08
;  EARLIER APPLICATION NUMBER: 60/051,916
;  EARLIER FILING DATE: 1997-07-08
;  EARLIER APPLICATION NUMBER: 60/051,930
;  EARLIER FILING DATE: 1997-07-08
;  EARLIER APPLICATION NUMBER: 60/051,918
;  EARLIER FILING DATE: 1997-07-08
;  EARLIER APPLICATION NUMBER: 60/051,920
;  EARLIER FILING DATE: 1997-07-08
;  EARLIER APPLICATION NUMBER: 60/052,733
;  EARLIER FILING DATE: 1997-07-08
;  EARLIER APPLICATION NUMBER: 60/052,795
;  EARLIER FILING DATE: 1997-07-08
;  EARLIER APPLICATION NUMBER: 60/051,919
;  EARLIER FILING DATE: 1997-07-08
;  EARLIER APPLICATION NUMBER: 60/051,928
;  EARLIER FILING DATE: 1997-07-08
;  EARLIER APPLICATION NUMBER: 60/055,722
;  EARLIER FILING DATE: 1997-08-18
;  EARLIER APPLICATION NUMBER: 60/055,723
;  EARLIER FILING DATE: 1997-08-18
;  EARLIER APPLICATION NUMBER: 60/055,948
;  EARLIER FILING DATE: 1997-08-18
;  EARLIER APPLICATION NUMBER: 60/055,949
;  EARLIER FILING DATE: 1997-08-18
;  EARLIER APPLICATION NUMBER: 60/055,953
;  EARLIER FILING DATE: 1997-08-18
;  EARLIER APPLICATION NUMBER: 60/055,950
;  EARLIER FILING DATE: 1997-08-18
;  EARLIER APPLICATION NUMBER: 60/055,947
;  EARLIER FILING DATE: 1997-08-18
;  EARLIER APPLICATION NUMBER: 60/055,964
;  EARLIER FILING DATE: 1997-08-18
;  EARLIER APPLICATION NUMBER: 60/056,360
;  EARLIER FILING DATE: 1997-08-18
;  EARLIER APPLICATION NUMBER: 60/055,684
;  EARLIER FILING DATE: 1997-08-18
;  EARLIER APPLICATION NUMBER: 60/055,984
;  EARLIER FILING DATE: 1997-08-18
;  EARLIER APPLICATION NUMBER: 60/055,954
;  EARLIER FILING DATE: 1997-08-18
;  EARLIER APPLICATION NUMBER: 60/058,785
;  EARLIER FILING DATE: 1997-09-12
;  EARLIER APPLICATION NUMBER: 60/058,664
;  EARLIER FILING DATE: 1997-09-12
;  EARLIER APPLICATION NUMBER: 60/058,660
;  EARLIER FILING DATE: 1997-09-12
;  EARLIER APPLICATION NUMBER: 60/058,661
;  EARLIER FILING DATE: 1997-09-12
;  NUMBER OF SEQ ID NOS: 672
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 190
;   LENGTH: 128
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (127)
;   OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (128)
;   OTHER INFORMATION: Xaa equals stop translation
US-09-227-357-190

  Query Match             4.5%;  Score 11;  DB 4;  Length 128;
  Best Local Similarity   100.0%;  Pred. No. 0.0074;
  Matches   11;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         45 GLPGRDGRDGR 55
              ||||||||||||
Db         59 GLPGRDGRDGR 69
```

```
RESULT 13
US-08-931-820-4
; Sequence 4, Application US/08931820
; Patent No. 6010863
;  GENERAL INFORMATION:
;    APPLICANT:
;    TITLE OF INVENTION:  Assay for collagen degradation
;    NUMBER OF SEQUENCES:  4
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25 (EPO)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/931,820
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  EP 96202596.1
;      FILING DATE:
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1057 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    HYPOTHETICAL:  NO
;    ORIGINAL SOURCE:
;      ORGANISM:  Homo sapiens
;      TISSUE TYPE:  Collagen type III
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  1055
;      OTHER INFORMATION:  /label= Modified
;      OTHER INFORMATION:  /note= "Ala may be Pro"
US-08-931-820-4

  Query Match             4.1%;  Score 10;  DB 3;  Length 1057;
  Best Local Similarity   100.0%;  Pred. No. 0.4;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         60 GAPGEKGEGG 69
              ||||||||||
Db        680 GAPGEKGEGG 689


RESULT 14
US-08-963-825-21
; Sequence 21, Application US/08963825
; Patent No. 6110689
;  GENERAL INFORMATION:
;    APPLICANT:  Qvist, Per
;    APPLICANT:  Bonde, Martin
;    TITLE OF INVENTION:  A Method for Assaying Collagen Fragments
;    TITLE OF INVENTION:  in Body Fluids, A Test Kit and Means for Carrying Out the
;    TITLE OF INVENTION:  Method and Use of the Method to Diagnose the Presence of
;    TITLE OF INVENTION:  Disorders Associated with the Metabolism of
;    NUMBER OF SEQUENCES:  21
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Darby & Darby PC
;      STREET:  805 Third Avenue
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10022
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/963,825
;      FILING DATE:
;      CLASSIFICATION:  436
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/187,319
;      FILING DATE:  21-JAN-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Gogoris, Adda C
;      REGISTRATION NUMBER:  29,714
;      REFERENCE/DOCKET NUMBER:  4305/08701
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  212-527-7700
;      TELEFAX:  212-753-6237
;      TELEX:  236687
;  INFORMATION FOR SEQ ID NO:  21:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1078 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    ORIGINAL SOURCE:
;      ORGANISM:  Homo sapiens
;    IMMEDIATE SOURCE:
;      CLONE:  COLLAGEN ALPHA 1 (III)
US-08-963-825-21

  Query Match             4.1%;  Score 10;  DB 3;  Length 1078;
  Best Local Similarity   100.0%;  Pred. No. 0.4;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         60 GAPGEKGEGG 69
              ||||||||||
Db        681 GAPGEKGEGG 690


RESULT 15
US-09-500-811-21
; Sequence 21, Application US/09500811
; Patent No. 6323314
;  GENERAL INFORMATION:
;    APPLICANT:  Qvist, Per
;    APPLICANT:  Bonde, Martin
;    TITLE OF INVENTION:  A Method for Assaying Collagen Fragments
;    TITLE OF INVENTION:  in Body Fluids, A Test Kit and Means for Carrying Out the
;    TITLE OF INVENTION:  Method and Use of the Method to Diagnose the Presence of
;    TITLE OF INVENTION:  Disorders Associated with the Metabolism of
;    NUMBER OF SEQUENCES:  21
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Darby & Darby PC
;      STREET:  805 Third Avenue
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10022
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/500,811
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/187,319
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Gogoris, Adda C
;      REGISTRATION NUMBER:  29,714
;      REFERENCE/DOCKET NUMBER:  4305/08701
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  212-527-7700
;      TELEFAX:  212-753-6237
;      TELEX:  236687
```

; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1078 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; IMMEDIATE SOURCE:
; CLONE: COLLAGEN ALPHA 1 (III)
US-09-500-811-21

Query Match 4.1%; Score 10; DB 4; Length 1078;
Best Local Similarity 100.0%; Pred. No. 0.4;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 60 GAPGEKGEGG 69
||||||||||
Db 681 GAPGEKGEGG 690

Search completed: June 18, 2003, 15:07:04
Job time : 27 secs

GenCore version 5.1.6


OM protein - protein search, using sw model

Run on: June 18, 2003, 15:05:58 ; Search time 49 Seconds
(without alignments)
536.617 Million cell updates/sec

Title: US-09-943-851A-42
Perfect score: 243
Sequence: 1 MRPLLVLLLLGLAAGSPPLD.....DSTFSGFLVYSDWHSSPVFA 243

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 417779 seqs, 108206813 residues

Word size : 0

Total number of hits satisfying chosen parameters: 417779

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 243 | 100.0 | 243 | 9 | US-09-944-413-42 | Sequence 42, Appl |
| 2 | 243 | 100.0 | 243 | 9 | US-09-944-403-42 | Sequence 42, Appl |
| 3 | 243 | 100.0 | 243 | 9 | US-09-944-896-42 | Sequence 42, Appl |
| 4 | 243 | 100.0 | 243 | 9 | US-09-944-944-42 | Sequence 42, Appl |
| 5 | 243 | 100.0 | 243 | 9 | US-09-944-907-42 | Sequence 42, Appl |
| 6 | 243 | 100.0 | 243 | 9 | US-09-944-929-42 | Sequence 42, Appl |
| 7 | 243 | 100.0 | 243 | 9 | US-10-028-072-362 | Sequence 362, App |
| 8 | 243 | 100.0 | 243 | 9 | US-10-121-049-362 | Sequence 362, App |
| 9 | 243 | 100.0 | 243 | 9 | US-10-123-904-362 | Sequence 362, App |
| 10 | 243 | 100.0 | 243 | 9 | US-10-140-470-362 | Sequence 362, App |
| 11 | 243 | 100.0 | 243 | 9 | US-09-796-753-68 | Sequence 68, Appl |
| 12 | 243 | 100.0 | 243 | 9 | US-10-175-746-362 | Sequence 362, App |
| 13 | 243 | 100.0 | 243 | 9 | US-10-176-918-362 | Sequence 362, App |
| 14 | 243 | 100.0 | 243 | 9 | US-10-176-921-362 | Sequence 362, App |
| 15 | 243 | 100.0 | 243 | 9 | US-10-137-865-362 | Sequence 362, App |
| 16 | 243 | 100.0 | 243 | 9 | US-10-140-474-362 | Sequence 362, App |
| 17 | 243 | 100.0 | 243 | 9 | US-10-142-431-362 | Sequence 362, App |
| 18 | 243 | 100.0 | 243 | 9 | US-10-143-114-362 | Sequence 362, App |
| 19 | 243 | 100.0 | 243 | 9 | US-10-140-002-362 | Sequence 362, App |
| 20 | 243 | 100.0 | 243 | 9 | US-10-142-419-362 | Sequence 362, App |
| 21 | 243 | 100.0 | 243 | 9 | US-10-123-262-362 | Sequence 362, App |
| 22 | 243 | 100.0 | 243 | 9 | US-10-142-423-362 | Sequence 362, App |
| 23 | 243 | 100.0 | 243 | 9 | US-10-121-050-362 | Sequence 362, App |
| 24 | 243 | 100.0 | 243 | 9 | US-10-141-755-362 | Sequence 362, App |
| 25 | 243 | 100.0 | 243 | 9 | US-10-143-032-362 | Sequence 362, App |
| 26 | 243 | 100.0 | 243 | 9 | US-10-123-108-362 | Sequence 362, App |
| 27 | 243 | 100.0 | 243 | 9 | US-10-123-236-362 | Sequence 362, App |
| 28 | 243 | 100.0 | 243 | 9 | US-10-123-261-362 | Sequence 362, App |
| 29 | 243 | 100.0 | 243 | 9 | US-10-140-921-362 | Sequence 362, App |
| 30 | 243 | 100.0 | 243 | 9 | US-10-140-928-362 | Sequence 362, App |
| 31 | 243 | 100.0 | 243 | 9 | US-10-121-045-362 | Sequence 362, App |
| 32 | 243 | 100.0 | 243 | 9 | US-10-123-292-362 | Sequence 362, App |
| 33 | 243 | 100.0 | 243 | 9 | US-10-123-903-362 | Sequence 362, App |
| 34 | 243 | 100.0 | 243 | 9 | US-10-124-819-362 | Sequence 362, App |
| 35 | 243 | 100.0 | 243 | 9 | US-10-124-822-362 | Sequence 362, App |
| 36 | 243 | 100.0 | 243 | 9 | US-10-140-925-362 | Sequence 362, App |
| 37 | 243 | 100.0 | 243 | 9 | US-10-160-498-362 | Sequence 362, App |
| 38 | 243 | 100.0 | 243 | 9 | US-09-944-884-42 | Sequence 42, Appl |
| 39 | 243 | 100.0 | 243 | 9 | US-10-121-041-362 | Sequence 362, App |
| 40 | 243 | 100.0 | 243 | 9 | US-10-121-043-362 | Sequence 362, App |
| 41 | 243 | 100.0 | 243 | 9 | US-10-121-047-362 | Sequence 362, App |
| 42 | 243 | 100.0 | 243 | 9 | US-10-123-215-362 | Sequence 362, App |
| 43 | 243 | 100.0 | 243 | 9 | US-10-123-902-362 | Sequence 362, App |
| 44 | 243 | 100.0 | 243 | 9 | US-10-123-908-362 | Sequence 362, App |
| 45 | 243 | 100.0 | 243 | 9 | US-10-123-909-362 | Sequence 362, App |

ALIGNMENTS

RESULT 1
US-09-944-413-42
; Sequence 42, Application US/09944413
; Patent No. US20020156004A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin
; APPLICANT: Botstein,David
; APPLICANT: Eaton,Dan
; APPLICANT: Ferrara,Napoleone
; APPLICANT: Filvaroff,Ellen
; APPLICANT: Gerritsen,Mary
; APPLICANT: Goddard,Audrey
; APPLICANT: Godowski,Paul
; APPLICANT: Grimaldi,Christopher
; APPLICANT: Gurney,Austin
; APPLICANT: Hillan,Kenneth
; APPLICANT: Kljavin,Ivar
; APPLICANT: Napier,Mary
; APPLICANT: Roy,Margaret
; APPLICANT: Tumas,Daniel
; APPLICANT: Wood,William
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P2548P1C1
; CURRENT APPLICATION NUMBER: US/09/944,413
; CURRENT FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 09/866,028
; PRIOR FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: 60/067,411
; PRIOR FILING DATE: December 3, 1997
; PRIOR APPLICATION NUMBER: 60/069,334
; PRIOR FILING DATE: December 11, 1997
; PRIOR APPLICATION NUMBER: 60/069335
; PRIOR FILING DATE: December 11, 1997
; PRIOR APPLICATION NUMBER: 60/069,278
; PRIOR FILING DATE: December 11, 1997
; PRIOR APPLICATION NUMBER: 60/069,425
; PRIOR FILING DATE: December 12, 1997
; PRIOR APPLICATION NUMBER: 60/069,696
; PRIOR FILING DATE: December 16, 1997
; PRIOR APPLICATION NUMBER: 60/069,694
; PRIOR FILING DATE: December 16, 1997

; PRIOR APPLICATION NUMBER: 60/069,702
; PRIOR FILING DATE: December 16, 1997
; PRIOR APPLICATION NUMBER: 60/069,870
; PRIOR FILING DATE: December 17, 1997
; PRIOR APPLICATION NUMBER: 60/069,873
; PRIOR FILING DATE: December 17, 1997
; PRIOR APPLICATION NUMBER: 60/068,017
; PRIOR FILING DATE: December 18, 1997
; PRIOR APPLICATION NUMBER: 60/070,440
; PRIOR FILING DATE: January 5, 1998
; PRIOR APPLICATION NUMBER: 60/074,086
; PRIOR FILING DATE: February 9, 1998
; PRIOR APPLICATION NUMBER: 60/074,092
; PRIOR FILING DATE: February 9, 1998
; PRIOR APPLICATION NUMBER: 60/075,945
; PRIOR FILING DATE: February 25, 1998
; PRIOR APPLICATION NUMBER: 60/112,850
; PRIOR FILING DATE: December 16, 1998
; PRIOR APPLICATION NUMBER: 60/113,296
; PRIOR FILING DATE: December 22, 1998
; PRIOR APPLICATION NUMBER: 60/146,222
; PRIOR FILING DATE: July 28, 1999
; PRIOR APPLICATION NUMBER: PCT/US98/19330
; PRIOR FILING DATE: September 16, 1998
; PRIOR APPLICATION NUMBER: PCT/US98/25108
; PRIOR FILING DATE: December 1, 1998
; PRIOR APPLICATION NUMBER: 09/216,021
; PRIOR FILING DATE: December 16, 1998
; PRIOR APPLICATION NUMBER: 09/218,517
; PRIOR FILING DATE: December 22, 1998
; PRIOR APPLICATION NUMBER: 09/254,311
; PRIOR FILING DATE: March 3, 1999
; PRIOR APPLICATION NUMBER: PCT/US99/12252
; PRIOR FILING DATE: June 22, 1999
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: September 15, 1999
; PRIOR APPLICATION NUMBER: PCT/US99/28409
; PRIOR FILING DATE: No. US20020156004Alember 30, 1999
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: No. US20020156004Alember 30, 1999
; PRIOR APPLICATION NUMBER: PCT/US99/28301
; PRIOR FILING DATE: Decemberl, 1999
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: December 16, 1999
; PRIOR APPLICATION NUMBER: PCT/US00/03565
; PRIOR FILING DATE: February 11, 2000
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: February 22, 2000
; PRIOR APPLICATION NUMBER: PCT/US00/05841
; PRIOR FILING DATE: March 2, 2000
; PRIOR APPLICATION NUMBER: PCT/US00/08439
; PRIOR FILING DATE: March 30, 2000
; PRIOR APPLICATION NUMBER: PCT/US00/14042
; PRIOR FILING DATE: May 22, 2000
; PRIOR APPLICATION NUMBER: PCT/US00/20710
; PRIOR FILING DATE: July 28, 2000
; PRIOR APPLICATION NUMBER: PCT/US00/32678
; PRIOR FILING DATE: December 1, 2000
; PRIOR APPLICATION NUMBER: PCT/US01/06520
; PRIOR FILING DATE: February 28, 2001
; NUMBER OF SEQ ID NOS: 120
; SEQ ID NO 42
; LENGTH: 243
; TYPE: PRT
; ORGANISM: Homo Sapien
US-09-944-413-42

Query Match 100.0%; Score 243; DB 9; Length 243;
Best Local Similarity 100.0%; Pred. No. 7.9e-202;
Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MRPLLVLLLLGLAAGSPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MRPLLVLLLLGLAAGSPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPG 60

Qy         61 APGEKGEGGRPGLPGPRGDPGPRGEAGPAGPTGPAGECSVPPRSAFSAKRSESRVPPPSD 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 APGEKGEGGRPGLPGPRGDPGPRGEAGPAGPTGPAGECSVPPRSAFSAKRSESRVPPPSD 120

Qy        121 APLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGESIASFFQ 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 APLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGESIASFFQ 180

Qy        181 FFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWHSSP 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 FFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWHSSP 240

Qy        241 VFA 243
              |||
Db        241 VFA 243
```

RESULT 2
US-09-944-403-42
; Sequence 42, Application US/09944403
; Patent No. US20020165143A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin
; APPLICANT: Botstein,David
; APPLICANT: Eaton,Dan
; APPLICANT: Ferrara,Napoleone
; APPLICANT: Filvaroff,Ellen
; APPLICANT: Gerritsen,Mary
; APPLICANT: Goddard,Audrey
; APPLICANT: Godowski,Paul
; APPLICANT: Grimaldi,Christopher
; APPLICANT: Gurney,Austin
; APPLICANT: Hillan,Kenneth
; APPLICANT: Kljavin,Ivar
; APPLICANT: Napier,Mary
; APPLICANT: Roy,Margaret
; APPLICANT: Tumas,Daniel
; APPLICANT: Wood,William
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P2548P1C1
; CURRENT APPLICATION NUMBER: US/09/944,403
; CURRENT FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 09/866,028
; PRIOR FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: 60/067,411
; PRIOR FILING DATE: December 3, 1997
; PRIOR APPLICATION NUMBER: 60/069,334
; PRIOR FILING DATE: December 11, 1997
; PRIOR APPLICATION NUMBER: 60/069335
; PRIOR FILING DATE: December 11, 1997
; PRIOR APPLICATION NUMBER: 60/069,278
; PRIOR FILING DATE: December 11, 1997
; PRIOR APPLICATION NUMBER: 60/069,425
; PRIOR FILING DATE: December 12, 1997
; PRIOR APPLICATION NUMBER: 60/069,696
; PRIOR FILING DATE: December 16, 1997
; PRIOR APPLICATION NUMBER: 60/069,694
; PRIOR FILING DATE: December 16, 1997
; PRIOR APPLICATION NUMBER: 60/069,702
; PRIOR FILING DATE: December 16, 1997
; PRIOR APPLICATION NUMBER: 60/069,870
; PRIOR FILING DATE: December 17, 1997
; PRIOR APPLICATION NUMBER: 60/069,873
; PRIOR FILING DATE: December 17, 1997
; PRIOR APPLICATION NUMBER: 60/068,017
; PRIOR FILING DATE: December 18, 1997
; PRIOR APPLICATION NUMBER: 60/070,440
; PRIOR FILING DATE: January 5, 1998
; PRIOR APPLICATION NUMBER: 60/074,086
; PRIOR FILING DATE: February 9, 1998

```
;  PRIOR APPLICATION NUMBER: 60/074,092
;  PRIOR FILING DATE: February 9, 1998
;  PRIOR APPLICATION NUMBER: 60/075,945
;  PRIOR FILING DATE: February 25, 1998
;  PRIOR APPLICATION NUMBER: 60/112,850
;  PRIOR FILING DATE: December 16, 1998
;  PRIOR APPLICATION NUMBER: 60/113,296
;  PRIOR FILING DATE: December 22, 1998
;  PRIOR APPLICATION NUMBER: 60/146,222
;  PRIOR FILING DATE: July 28, 1999
;  PRIOR APPLICATION NUMBER: PCT/US98/19330
;  PRIOR FILING DATE: September 16, 1998
;  PRIOR APPLICATION NUMBER: PCT/US98/25108
;  PRIOR FILING DATE: December 1, 1998
;  PRIOR APPLICATION NUMBER: 09/216,021
;  PRIOR FILING DATE: December 16, 1998
;  PRIOR APPLICATION NUMBER: 09/218,517
;  PRIOR FILING DATE: December 22, 1998
;  PRIOR APPLICATION NUMBER: 09/254,311
;  PRIOR FILING DATE: March 3, 1999
;  PRIOR APPLICATION NUMBER: PCT/US99/12252
;  PRIOR FILING DATE: June 22, 1999
;  PRIOR APPLICATION NUMBER: PCT/US99/21090
;  PRIOR FILING DATE: September 15, 1999
;  PRIOR APPLICATION NUMBER: PCT/US99/28409
;  PRIOR FILING DATE: No. US20020165143Alember 30, 1999
;  PRIOR APPLICATION NUMBER: PCT/US99/28313
;  PRIOR FILING DATE: No. US20020165143Alember 30, 1999
;  PRIOR APPLICATION NUMBER: PCT/US99/28301
;  PRIOR FILING DATE: Decemberl, 1999
;  PRIOR APPLICATION NUMBER: PCT/US99/30095
;  PRIOR FILING DATE: December 16, 1999
;  PRIOR APPLICATION NUMBER: PCT/US00/03565
;  PRIOR FILING DATE: February 11, 2000
;  PRIOR APPLICATION NUMBER: PCT/US00/04414
;  PRIOR FILING DATE: February 22, 2000
;  PRIOR APPLICATION NUMBER: PCT/US00/05841
;  PRIOR FILING DATE: March 2, 2000
;  PRIOR APPLICATION NUMBER: PCT/US00/08439
;  PRIOR FILING DATE: March 30, 2000
;  PRIOR APPLICATION NUMBER: PCT/US00/14042
;  PRIOR FILING DATE: May 22, 2000
;  PRIOR APPLICATION NUMBER: PCT/US00/20710
;  PRIOR FILING DATE: July 28, 2000
;  PRIOR APPLICATION NUMBER: PCT/US00/32678
;  PRIOR FILING DATE: December 1, 2000
;  PRIOR APPLICATION NUMBER: PCT/US01/06520
;  PRIOR FILING DATE: February 28, 2001
;  NUMBER OF SEQ ID NOS: 120
; SEQ ID NO 42
;   LENGTH: 243
;   TYPE: PRT
;   ORGANISM: Homo Sapien
US-09-944-403-42

  Query Match           100.0%;  Score 243;  DB 9;  Length 243;
  Best Local Similarity  100.0%;  Pred. No. 7.9e-202;
  Matches  243;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MRPLLVLLLLGLAAGSPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MRPLLVLLLLGLAAGSPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPG 60

Qy         61 APGEKGEGGRPGLPGPRGDPGPRGEAGPAGPTGPAGECSVPPRSAFSAKRSESRVPPPSD 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 APGEKGEGGRPGLPGPRGDPGPRGEAGPAGPTGPAGECSVPPRSAFSAKRSESRVPPPSD 120

Qy        121 APLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGESIASFFQ 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 APLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGESIASFFQ 180

Qy        181 FFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWHSSP 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 FFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWHSSP 240

Qy        241 VFA 243
              |||
Db        241 VFA 243


RESULT 3
US-09-944-896-42
; Sequence 42, Application US/09944896
; Patent No. US20020168715A1
; GENERAL INFORMATION:
;  APPLICANT: Baker, Kevin
;  APPLICANT:  Botstein,David
;  APPLICANT:  Eaton,Dan
;  APPLICANT:  Ferrara,Napoleone
;  APPLICANT:  Filvaroff,Ellen
;  APPLICANT:  Gerritsen,Mary
;  APPLICANT:  Goddard,Audrey
;  APPLICANT:  Godowski,Paul
;  APPLICANT:  Grimaldi,Christopher
;  APPLICANT:  Gurney,Austin
;  APPLICANT:  Hillan,Kenneth
;  APPLICANT:  Kljavin,Ivar
;  APPLICANT:  Napier,Mary
;  APPLICANT:  Roy,Margaret
;  APPLICANT:  Tumas,Daniel
;  APPLICANT:  Wood,William
;  TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;  TITLE OF INVENTION:  ACIDS ENCODING THE SAME
;  FILE REFERENCE: P2548P1C1
;  CURRENT APPLICATION NUMBER: US/09/944,896
;  CURRENT FILING DATE:  2001-08-31
;  PRIOR APPLICATION NUMBER: 09/866,028
;  PRIOR FILING DATE: 2001-05-25
;  PRIOR APPLICATION NUMBER: 60/069,334
;  PRIOR FILING DATE: December 11, 1997
;  PRIOR APPLICATION NUMBER: 60/069335
;  PRIOR FILING DATE: December 11, 1997
;  PRIOR APPLICATION NUMBER: 60/069,278
;  PRIOR FILING DATE: December 11, 1997
;  PRIOR APPLICATION NUMBER: 60/069,425
;  PRIOR FILING DATE: December 12, 1997
;  PRIOR APPLICATION NUMBER: 60/069,696
;  PRIOR FILING DATE: December 16, 1997
;  PRIOR APPLICATION NUMBER: 60/069,694
;  PRIOR FILING DATE: December 16, 1997
;  PRIOR APPLICATION NUMBER: 60/069,702
;  PRIOR FILING DATE: December 16, 1997
;  PRIOR APPLICATION NUMBER: 60/069,870
;  PRIOR FILING DATE: December 17, 1997
;  PRIOR APPLICATION NUMBER: 60/069,873
;  PRIOR FILING DATE: December 17, 1997
;  PRIOR APPLICATION NUMBER: 60/068,017
;  PRIOR FILING DATE: December 18, 1997
;  PRIOR APPLICATION NUMBER: 60/070,440
;  PRIOR FILING DATE: January 5, 1998
;  PRIOR APPLICATION NUMBER: 60/074,086
;  PRIOR FILING DATE: February 9, 1998
;  PRIOR APPLICATION NUMBER: 60/074,092
;  PRIOR FILING DATE: February 9, 1998
;  PRIOR APPLICATION NUMBER: 60/075,945
;  PRIOR FILING DATE: February 25, 1998
;  PRIOR APPLICATION NUMBER: 60/112,850
;  PRIOR FILING DATE: December 16, 1998
;  PRIOR APPLICATION NUMBER: 60/113,296
;  PRIOR FILING DATE: December 22, 1998
;  PRIOR APPLICATION NUMBER: 60/146,222
;  PRIOR FILING DATE: July 28, 1999
;  PRIOR APPLICATION NUMBER: PCT/US98/19330
;  PRIOR FILING DATE: September 16, 1998
;  PRIOR APPLICATION NUMBER: PCT/US98/25108
;  PRIOR FILING DATE: December 1, 1998
```

```
;  PRIOR APPLICATION NUMBER: 09/216,021
;  PRIOR FILING DATE: December 16, 1998
;  PRIOR APPLICATION NUMBER: 09/218,517
;  PRIOR FILING DATE: December 22, 1998
;  PRIOR APPLICATION NUMBER: 09/254,311
;  PRIOR FILING DATE: March 3, 1999
;  PRIOR APPLICATION NUMBER: PCT/US99/12252
;  PRIOR FILING DATE: June 22, 1999
;  PRIOR APPLICATION NUMBER: PCT/US99/21090
;  PRIOR FILING DATE: September 15, 1999
;  PRIOR APPLICATION NUMBER: PCT/US99/28409
;  PRIOR FILING DATE: No. US20020168715Alember 30, 1999
;  PRIOR APPLICATION NUMBER: PCT/US99/28313
;  PRIOR FILING DATE: No. US20020168715Alember 30, 1999
;  PRIOR APPLICATION NUMBER: PCT/US99/28301
;  PRIOR FILING DATE: Decemberl, 1999
;  PRIOR APPLICATION NUMBER: PCT/US99/30095
;  PRIOR FILING DATE: December 16, 1999
;  PRIOR APPLICATION NUMBER: PCT/US00/03565
;  PRIOR FILING DATE: February 11, 2000
;  PRIOR APPLICATION NUMBER: PCT/US00/04414
;  PRIOR FILING DATE: February 22, 2000
;  PRIOR APPLICATION NUMBER: PCT/US00/05841
;  PRIOR FILING DATE: March 2, 2000
;  PRIOR APPLICATION NUMBER: PCT/US00/08439
;  PRIOR FILING DATE: March 30, 2000
;  PRIOR APPLICATION NUMBER: PCT/US00/14042
;  PRIOR FILING DATE: May 22, 2000
;  PRIOR APPLICATION NUMBER: PCT/US00/20710
;  PRIOR FILING DATE: July 28, 2000
;  PRIOR APPLICATION NUMBER: PCT/US00/32678
;  PRIOR FILING DATE: December 1, 2000
;  PRIOR APPLICATION NUMBER: PCT/US01/06520
;  PRIOR FILING DATE: February 28, 2001
;  NUMBER OF SEQ ID NOS: 120
; SEQ ID NO 42
;    LENGTH: 243
;    TYPE: PRT
;    ORGANISM: Homo Sapien
US-09-944-896-42

  Query Match             100.0%;  Score 243;  DB 9;  Length 243;
  Best Local Similarity   100.0%;  Pred. No. 7.9e-202;
  Matches  243;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MRPLLVLLLLGLAAGSPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MRPLLVLLLLGLAAGSPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPG 60

Qy         61 APGEKGEGGRPGLPGPRGDPGPRGEAGPAGPTGPAGECSVPPRSAFSAKRSESRVPPPSD 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 APGEKGEGGRPGLPGPRGDPGPRGEAGPAGPTGPAGECSVPPRSAFSAKRSESRVPPPSD 120

Qy        121 APLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGESIASFFQ 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 APLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGESIASFFQ 180

Qy        181 FFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWHSSP 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 FFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWHSSP 240

Qy        241 VFA 243
              |||
Db        241 VFA 243


RESULT 4
US-09-944-944-42
; Sequence 42, Application US/09944944
; Patent No. US20020173463A1
; GENERAL INFORMATION:
;  APPLICANT: Baker, Kevin
;  APPLICANT:  Botstein,David
;  APPLICANT:  Eaton,Dan
;  APPLICANT:  Ferrara,Napoleone
;  APPLICANT:  Filvaroff,Ellen
;  APPLICANT:  Gerritsen,Mary
;  APPLICANT:  Goddard,Audrey
;  APPLICANT:  Godowski,Paul
;  APPLICANT:  Grimaldi,Christopher
;  APPLICANT:  Gurney,Austin
;  APPLICANT:  Hillan,Kenneth
;  APPLICANT:  Kljavin,Ivar
;  APPLICANT:  Napier,Mary
;  APPLICANT:  Roy,Margaret
;  APPLICANT:  Tumas,Daniel
;  APPLICANT:  Wood,William
;  TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;  TITLE OF INVENTION:  ACIDS ENCODING THE SAME
;  FILE REFERENCE: P2548P1C1
;  CURRENT APPLICATION NUMBER: US/09/944,944
;  CURRENT FILING DATE:  2001-09-26
;  PRIOR APPLICATION NUMBER: 09/866,028
;  PRIOR FILING DATE: 2001-05-25
;  PRIOR APPLICATION NUMBER: 60/067,411
;  PRIOR FILING DATE: December 3, 1997
;  PRIOR APPLICATION NUMBER: 60/069,334
;  PRIOR FILING DATE: December 11, 1997
;  PRIOR APPLICATION NUMBER: 60/069335
;  PRIOR FILING DATE: December 11, 1997
;  PRIOR APPLICATION NUMBER: 60/069,278
;  PRIOR FILING DATE: December 11, 1997
;  PRIOR APPLICATION NUMBER: 60/069,425
;  PRIOR FILING DATE: December 12, 1997
;  PRIOR APPLICATION NUMBER: 60/069,696
;  PRIOR FILING DATE: December 16, 1997
;  PRIOR APPLICATION NUMBER: 60/069,694
;  PRIOR FILING DATE: December 16, 1997
;  PRIOR APPLICATION NUMBER: 60/069,702
;  PRIOR FILING DATE: December 16, 1997
;  PRIOR APPLICATION NUMBER: 60/069,870
;  PRIOR FILING DATE: December 17, 1997
;  PRIOR APPLICATION NUMBER: 60/069,873
;  PRIOR FILING DATE: December 17, 1997
;  PRIOR APPLICATION NUMBER: 60/068,017
;  PRIOR FILING DATE: December 18, 1997
;  PRIOR APPLICATION NUMBER: 60/070,440
;  PRIOR FILING DATE: January 5, 1998
;  PRIOR APPLICATION NUMBER: 60/074,086
;  PRIOR FILING DATE: February 9, 1998
;  PRIOR APPLICATION NUMBER: 60/074,092
;  PRIOR FILING DATE: February 9, 1998
;  PRIOR APPLICATION NUMBER: 60/075,945
;  PRIOR FILING DATE: February 25, 1998
;  PRIOR APPLICATION NUMBER: 60/112,850
;  PRIOR FILING DATE: December 16, 1998
;  PRIOR APPLICATION NUMBER: 60/113,296
;  PRIOR FILING DATE: December 22, 1998
;  PRIOR APPLICATION NUMBER: 60/146,222
;  PRIOR FILING DATE: July 28, 1999
;  PRIOR APPLICATION NUMBER: PCT/US98/19330
;  PRIOR FILING DATE: September 16, 1998
;  PRIOR APPLICATION NUMBER: PCT/US98/25108
;  PRIOR FILING DATE: December 1, 1998
;  PRIOR APPLICATION NUMBER: 09/216,021
;  PRIOR FILING DATE: December 16, 1998
;  PRIOR APPLICATION NUMBER: 09/218,517
;  PRIOR FILING DATE: December 22, 1998
;  PRIOR APPLICATION NUMBER: 09/254,311
;  PRIOR FILING DATE: March 3, 1999
;  PRIOR APPLICATION NUMBER: PCT/US99/12252
;  PRIOR FILING DATE: June 22, 1999
;  PRIOR APPLICATION NUMBER: PCT/US99/21090
;  PRIOR FILING DATE: September 15, 1999
;  PRIOR APPLICATION NUMBER: PCT/US99/28409
;  PRIOR FILING DATE: No. US20020173463Alember 30, 1999
```

```
;  PRIOR APPLICATION NUMBER: PCT/US99/28313
;  PRIOR FILING DATE: No. US20020173463A1ember 30, 1999
;  PRIOR APPLICATION NUMBER: PCT/US99/28301
;  PRIOR FILING DATE: December1, 1999
;  PRIOR APPLICATION NUMBER: PCT/US99/30095
;  PRIOR FILING DATE: December 16, 1999
;  PRIOR APPLICATION NUMBER: PCT/US00/03565
;  PRIOR FILING DATE: February 11, 2000
;  PRIOR APPLICATION NUMBER: PCT/US00/04414
;  PRIOR FILING DATE: February 22, 2000
;  PRIOR APPLICATION NUMBER: PCT/US00/05841
;  PRIOR FILING DATE: March 2, 2000
;  PRIOR APPLICATION NUMBER: PCT/US00/08439
;  PRIOR FILING DATE: March 30, 2000
;  PRIOR APPLICATION NUMBER: PCT/US00/14042
;  PRIOR FILING DATE: May 22, 2000
;  PRIOR APPLICATION NUMBER: PCT/US00/20710
;  PRIOR FILING DATE: July 28, 2000
;  PRIOR APPLICATION NUMBER: PCT/US00/32678
;  PRIOR FILING DATE: December 1, 2000
;  PRIOR APPLICATION NUMBER: PCT/US01/06520
;  PRIOR FILING DATE: February 28, 2001
;  NUMBER OF SEQ ID NOS: 120
; SEQ ID NO 42
;   LENGTH: 243
;   TYPE: PRT
;   ORGANISM: Homo Sapien
US-09-944-944-42

  Query Match           100.0%;  Score 243;  DB 9;  Length 243;
  Best Local Similarity  100.0%;  Pred. No. 7.9e-202;
  Matches  243;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MRPLLVLLLLGLAAGSPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MRPLLVLLLLGLAAGSPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPG 60

Qy         61 APGEKGEGGRPGLPGPRGDPGPRGEAGPAGPTGPAGECSVPPRSAFSAKRSESRVPPPSD 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 APGEKGEGGRPGLPGPRGDPGPRGEAGPAGPTGPAGECSVPPRSAFSAKRSESRVPPPSD 120

Qy        121 APLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGESIASFFQ 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 APLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGESIASFFQ 180

Qy        181 FFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWHSSP 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 FFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWHSSP 240

Qy        241 VFA 243
              |||
Db        241 VFA 243


RESULT 5
US-09-944-907-42
; Sequence 42, Application US/09944907
; Publication No. US20020198147A1
; GENERAL INFORMATION:
;  APPLICANT: Baker, Kevin
;  APPLICANT:  Botstein,David
;  APPLICANT:  Eaton,Dan
;  APPLICANT:  Ferrara,Napoleone
;  APPLICANT:  Filvaroff,Ellen
;  APPLICANT:  Gerritsen,Mary
;  APPLICANT:  Goddard,Audrey
;  APPLICANT:  Godowski,Paul
;  APPLICANT:  Grimaldi,Christopher
;  APPLICANT:  Gurney,Austin
;  APPLICANT:  Hillan,Kenneth
;  APPLICANT:  Kljavin,Ivar
;  APPLICANT:  Napier,Mary
;  APPLICANT:  Roy,Margaret
;  APPLICANT:  Tumas,Daniel
;  APPLICANT:  Wood,William
;  TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;  TITLE OF INVENTION:  ACIDS ENCODING THE SAME
;  FILE REFERENCE: P2548P1C1
;  CURRENT APPLICATION NUMBER: US/09/944,907
;  CURRENT FILING DATE:  2001-08-31
;  PRIOR APPLICATION NUMBER: 09/866,028
;  PRIOR FILING DATE: 2001-05-25
;  NUMBER OF SEQ ID NOS: 120
; SEQ ID NO 42
;   LENGTH: 243
;   TYPE: PRT
;   ORGANISM: Homo Sapien
US-09-944-907-42

  Query Match           100.0%;  Score 243;  DB 9;  Length 243;
  Best Local Similarity  100.0%;  Pred. No. 7.9e-202;
  Matches  243;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MRPLLVLLLLGLAAGSPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MRPLLVLLLLGLAAGSPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPG 60

Qy         61 APGEKGEGGRPGLPGPRGDPGPRGEAGPAGPTGPAGECSVPPRSAFSAKRSESRVPPPSD 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 APGEKGEGGRPGLPGPRGDPGPRGEAGPAGPTGPAGECSVPPRSAFSAKRSESRVPPPSD 120

Qy        121 APLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGESIASFFQ 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 APLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGESIASFFQ 180

Qy        181 FFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWHSSP 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 FFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWHSSP 240

Qy        241 VFA 243
              |||
Db        241 VFA 243


RESULT 6
US-09-944-929-42
; Sequence 42, Application US/09944929
; Publication No. US20020197612A1
; GENERAL INFORMATION:
;  APPLICANT: Baker, Kevin
;  APPLICANT:  Botstein,David
;  APPLICANT:  Eaton,Dan
;  APPLICANT:  Ferrara,Napoleone
;  APPLICANT:  Filvaroff,Ellen
;  APPLICANT:  Gerritsen,Mary
;  APPLICANT:  Goddard,Audrey
;  APPLICANT:  Godowski,Paul
;  APPLICANT:  Grimaldi,Christopher
;  APPLICANT:  Gurney,Austin
;  APPLICANT:  Hillan,Kenneth
;  APPLICANT:  Kljavin,Ivar
;  APPLICANT:  Napier,Mary
;  APPLICANT:  Roy,Margaret
;  APPLICANT:  Tumas,Daniel
;  APPLICANT:  Wood,William
;  TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;  TITLE OF INVENTION:  ACIDS ENCODING THE SAME
;  FILE REFERENCE: P2548P1C1
;  CURRENT APPLICATION NUMBER: US/09/944,929
;  CURRENT FILING DATE:  2001-08-31
;  PRIOR APPLICATION NUMBER: 09/866,028
;  PRIOR FILING DATE: 2001-05-25
;  NUMBER OF SEQ ID NOS: 120
; SEQ ID NO 42
;   LENGTH: 243
;   TYPE: PRT
```

```
;  ORGANISM: Homo Sapien
US-09-944-929-42

  Query Match           100.0%;  Score 243;  DB 9;  Length 243;
  Best Local Similarity  100.0%;  Pred. No. 7.9e-202;
  Matches  243;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MRPLLVLLLLGLAAGSPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MRPLLVLLLLGLAAGSPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPG 60

Qy         61 APGEKGEGGRPGLPGPRGDPGPRGEAGPAGPTGPAGECSVPPRSAFSAKRSESRVPPPSD 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 APGEKGEGGRPGLPGPRGDPGPRGEAGPAGPTGPAGECSVPPRSAFSAKRSESRVPPPSD 120

Qy        121 APLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGESIASFFQ 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 APLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGESIASFFQ 180

Qy        181 FFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWHSSP 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 FFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWHSSP 240

Qy        241 VFA 243
              |||
Db        241 VFA 243


RESULT 7
US-10-028-072-362
; Sequence 362, Application US/10028072
; Publication No. US20030004311A1
; GENERAL INFORMATION:
;  APPLICANT: Baker,Kevin P.
;  APPLICANT:  Beresini,Maureen
;  APPLICANT:  DeForge,Laura
;  APPLICANT:  Desnoyers,Luc
;  APPLICANT:  Filvaroff,Ellen
;  APPLICANT:  Gao,Wei-Qiang
;  APPLICANT:  Gerritsen,Mary E.
;  APPLICANT:  Goddard,Audrey
;  APPLICANT:  Godowski,Paul J.
;  APPLICANT:  Gurney,Austin L.
;  APPLICANT:  Sherwood,Steven
;  APPLICANT:  Smith,Victoria
;  APPLICANT:  Stewart,Timothy A.
;  APPLICANT:  Tumas,Daniel
;  APPLICANT:  Watanabe,Colin K
;  APPLICANT:  Wood,William
;  APPLICANT:  Zhang
;  TITLE OF INVENTION:
;  FILE REFERENCE:
;  CURRENT APPLICATION NUMBER: US/10/028,072
;  CURRENT FILING DATE:  2001-12-19
;  PRIOR APPLICATION NUMBER: 60/049911
;  PRIOR FILING DATE: 1997-06-18
;  PRIOR APPLICATION NUMBER: 60/056974
;  PRIOR FILING DATE: 1997-08-26
;  PRIOR APPLICATION NUMBER: 60/059113
;  PRIOR FILING DATE: 1997-09-17
;  PRIOR APPLICATION NUMBER: 60/059115
;  PRIOR FILING DATE: 1997-09-17
;  PRIOR APPLICATION NUMBER: 60/059117
;  PRIOR FILING DATE: 1997-09-17
;  PRIOR APPLICATION NUMBER: 60/059122
;  PRIOR FILING DATE: 1997-09-17
;  PRIOR APPLICATION NUMBER: 60/059184
;  PRIOR FILING DATE: 1997-09-17
;  PRIOR APPLICATION NUMBER: 60/059263
;  PRIOR FILING DATE: 1997-09-18
;  PRIOR APPLICATION NUMBER: 60/059352
;  PRIOR FILING DATE: 1997-09-19
;  PRIOR APPLICATION NUMBER: 60/059588
;  PRIOR FILING DATE: 1997-09-19
;  PRIOR APPLICATION NUMBER: 60/059836
;  PRIOR FILING DATE: 1997-09-24
;  PRIOR APPLICATION NUMBER: 60/062250
;  PRIOR FILING DATE: 1997-10-17
;  PRIOR APPLICATION NUMBER: 60/062285
;  PRIOR FILING DATE: 1997-10-17
;  PRIOR APPLICATION NUMBER: 60/062287
;  PRIOR FILING DATE: 1997-10-17
;  PRIOR APPLICATION NUMBER: 60/062814
;  PRIOR FILING DATE: 1997-10-24
;  PRIOR APPLICATION NUMBER: 60/062816
;  PRIOR FILING DATE: 1997-10-24
;  PRIOR APPLICATION NUMBER: 60/063045
;  PRIOR FILING DATE: 1997-10-24
;  PRIOR APPLICATION NUMBER: 60/063082
;  PRIOR FILING DATE: 1997-10-31
;  PRIOR APPLICATION NUMBER: 60/063127
;  PRIOR FILING DATE: 1997-10-24
;  PRIOR APPLICATION NUMBER: 60/063327
;  PRIOR FILING DATE: 1997-10-27
;  PRIOR APPLICATION NUMBER: 60/063329
;  PRIOR FILING DATE: 1997-10-27
;  PRIOR APPLICATION NUMBER: 60/063550
;  PRIOR FILING DATE: 1997-10-28
;  PRIOR APPLICATION NUMBER: 60/063561
;  PRIOR FILING DATE: 1997-10-28
;  PRIOR APPLICATION NUMBER: 60/063704
;  PRIOR FILING DATE: 1997-10-29
;  PRIOR APPLICATION NUMBER: 60/063733
;  PRIOR FILING DATE: 1997-10-29
;  PRIOR APPLICATION NUMBER: 60/063735
;  PRIOR FILING DATE: 1997-10-29
;  PRIOR APPLICATION NUMBER: 60/063738
;  PRIOR FILING DATE: 1997-10-29
;  PRIOR APPLICATION NUMBER: 60/063755
;  PRIOR FILING DATE: 1997-10-17
;  PRIOR APPLICATION NUMBER: 60/064248
;  PRIOR FILING DATE: 1997-11-03
;  PRIOR APPLICATION NUMBER: 60/064809
;  PRIOR FILING DATE: 1997-11-07
;  PRIOR APPLICATION NUMBER: 60/065186
;  PRIOR FILING DATE: 1997-11-12
;  PRIOR APPLICATION NUMBER: 60/065846
;  PRIOR FILING DATE: 1997-11-17
;  PRIOR APPLICATION NUMBER: 60/066364
;  PRIOR FILING DATE: 1997-11-21
;  PRIOR APPLICATION NUMBER: 60/066453
;  PRIOR FILING DATE: 1997-11-24
;  PRIOR APPLICATION NUMBER: 60/066511
;  PRIOR FILING DATE: 1997-11-24
;  PRIOR APPLICATION NUMBER: 60/066770
;  PRIOR FILING DATE: 1997-11-24
;  PRIOR APPLICATION NUMBER: 60/069212
;  PRIOR FILING DATE: 1997-12-11
;  PRIOR APPLICATION NUMBER: 60/069278
;  PRIOR FILING DATE: 1997-12-11
;  PRIOR APPLICATION NUMBER: 60/069334
;  PRIOR FILING DATE: 1997-12-11
;  PRIOR APPLICATION NUMBER: 60/069694
;  PRIOR FILING DATE: 1997-12-16
;  PRIOR APPLICATION NUMBER: 60/072320
;  PRIOR FILING DATE: 1998-01-23
;  PRIOR APPLICATION NUMBER: 60/073612
;  PRIOR FILING DATE: 1998-02-04
;  PRIOR APPLICATION NUMBER: 60/074086
;  PRIOR FILING DATE: 1998-02-09
;  PRIOR APPLICATION NUMBER: 60/074092
;  PRIOR FILING DATE: 1998-02-09
;  PRIOR APPLICATION NUMBER: 60/077791
;  PRIOR FILING DATE: 1998-03-12
;  PRIOR APPLICATION NUMBER: 60/078910
;  PRIOR FILING DATE: 1998-03-20
```

; PRIOR APPLICATION NUMBER: 60/079294
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: 60/079663
; PRIOR FILING DATE: 1998-02-27
; PRIOR APPLICATION NUMBER: 60/079728
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/080165
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/081203
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: 60/081229
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: 60/081695
; PRIOR FILING DATE: 1998-04-14
; PRIOR APPLICATION NUMBER: 60/081817
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081818
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/082999
; PRIOR FILING DATE: 1998-04-24
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/083545
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084627
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084637
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/085149
; PRIOR FILING DATE: 1998-05-12
; PRIOR APPLICATION NUMBER: 60/085323
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085338
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085339
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085579
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085697
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085704
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/086414
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/086430
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/087106
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: 60/088026
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088730
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088741
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088810
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088858
; PRIOR FILING DATE: 19/98-06-11
; PRIOR APPLICATION NUMBER: 60/089532
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089599
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089907
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089947
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/090349
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090429
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090445
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090538
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090863
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/091360
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091519
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091982
; PRIOR FILING DATE: 1998-07-07

Query Match 100.0%; Score 243; DB 9; Length 243;
Best Local Similarity 100.0%; Pred. No. 7.9e-202;
Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MRPLLVLLLLGLAAGSPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MRPLLVLLLLGLAAGSPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPG 60

Qy         61 APGEKGEGGRPGLPGPRGDPGPRGEAGPAGPTGPAGECSVPPRSAFSAKRSESRVPPPSD 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 APGEKGEGGRPGLPGPRGDPGPRGEAGPAGPTGPAGECSVPPRSAFSAKRSESRVPPPSD 120

Qy        121 APLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGESIASFFQ 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 APLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGESIASFFQ 180

Qy        181 FFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWHSSP 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 FFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWHSSP 240

Qy        241 VFA 243
              |||
Db        241 VFA 243
```

RESULT 8
US-10-121-049-362
; Sequence 362, Application US/10121049
; Publication No. US20030022239A1
; GENERAL INFORMATION:
; APPLICANT: Baker,Kevin P.
; APPLICANT: Beresini,Maureen
; APPLICANT: DeForge,Laura
; APPLICANT: Desnoyers,Luc
; APPLICANT: Filvaroff,Ellen
; APPLICANT: Gao,Wei-Qiang
; APPLICANT: Gerritsen,Mary E.
; APPLICANT: Goddard,Audrey
; APPLICANT: Godowski,Paul J.
; APPLICANT: Gurney,Austin L.
; APPLICANT: Sherwood,Steven
; APPLICANT: Smith,Victoria
; APPLICANT: Stewart,Timothy A.
; APPLICANT: Tumas,Daniel
; APPLICANT: Watanabe,Colin K
; APPLICANT: Wood,William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3330R1C17
; CURRENT APPLICATION NUMBER: US/10/121,049
; CURRENT FILING DATE: 2002-04-12
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 362
; LENGTH: 243
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-121-049-362

Query Match 100.0%; Score 243; DB 9; Length 243;

```
  Best Local Similarity   100.0%;  Pred. No. 7.9e-202;
  Matches  243;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MRPLLVLLLLGLAAGSPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MRPLLVLLLLGLAAGSPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPG 60

Qy         61 APGEKGEGGRPGLPGPRGDPGPRGEAGPAGPTGPAGECSVPPRSAFSAKRSESRVPPPSD 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 APGEKGEGGRPGLPGPRGDPGPRGEAGPAGPTGPAGECSVPPRSAFSAKRSESRVPPPSD 120

Qy        121 APLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGESIASFFQ 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 APLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGESIASFFQ 180

Qy        181 FFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWHSSP 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 FFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWHSSP 240

Qy        241 VFA 243
              |||
Db        241 VFA 243


RESULT 9
US-10-123-904-362
; Sequence 362, Application US/10123904
; Publication No. US20030022328A1
; GENERAL INFORMATION:
;  APPLICANT: Baker,Kevin P.
;  APPLICANT:  Beresini,Maureen
;  APPLICANT:  DeForge,Laura
;  APPLICANT:  Desnoyers,Luc
;  APPLICANT:  Filvaroff,Ellen
;  APPLICANT:  Gao,Wei-Qiang
;  APPLICANT:  Gerritsen,Mary E.
;  APPLICANT:  Goddard,Audrey
;  APPLICANT:  Godowski,Paul J.
;  APPLICANT:  Gurney,Austin L.
;  APPLICANT:  Sherwood,Steven
;  APPLICANT:  Smith,Victoria
;  APPLICANT:  Stewart,Timothy A.
;  APPLICANT:  Tumas,Daniel
;  APPLICANT:  Watanabe,Colin K
;  APPLICANT:  Wood,William
;  APPLICANT:  Zhang, Zemin
;  TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;  TITLE OF INVENTION:  ACIDS ENCODING THE SAME
;  FILE REFERENCE: P3330R1C54
;  CURRENT APPLICATION NUMBER: US/10/123,904
;  CURRENT FILING DATE:  2002-04-16
;  Prior Application removed - See File Wrapper or Palm
;  NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 362
;   LENGTH: 243
;   TYPE: PRT
;   ORGANISM: Homo Sapien
US-10-123-904-362


  Query Match             100.0%;  Score 243;  DB 9;  Length 243;
  Best Local Similarity   100.0%;  Pred. No. 7.9e-202;
  Matches  243;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MRPLLVLLLLGLAAGSPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MRPLLVLLLLGLAAGSPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPG 60

Qy         61 APGEKGEGGRPGLPGPRGDPGPRGEAGPAGPTGPAGECSVPPRSAFSAKRSESRVPPPSD 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 APGEKGEGGRPGLPGPRGDPGPRGEAGPAGPTGPAGECSVPPRSAFSAKRSESRVPPPSD 120

Qy        121 APLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGESIASFFQ 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 APLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGESIASFFQ 180

Qy        181 FFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWHSSP 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 FFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWHSSP 240

Qy        241 VFA 243
              |||
Db        241 VFA 243


RESULT 10
US-10-140-470-362
; Sequence 362, Application US/10140470
; Publication No. US20030022331A1
; GENERAL INFORMATION:
;  APPLICANT: Baker,Kevin P.
;  APPLICANT:  Beresini,Maureen
;  APPLICANT:  DeForge,Laura
;  APPLICANT:  Desnoyers,Luc
;  APPLICANT:  Filvaroff,Ellen
;  APPLICANT:  Gao,Wei-Qiang
;  APPLICANT:  Gerritsen,Mary E.
;  APPLICANT:  Goddard,Audrey
;  APPLICANT:  Godowski,Paul J.
;  APPLICANT:  Gurney,Austin L.
;  APPLICANT:  Sherwood,Steven
;  APPLICANT:  Smith,Victoria
;  APPLICANT:  Stewart,Timothy A.
;  APPLICANT:  Tumas,Daniel
;  APPLICANT:  Watanabe,Colin K
;  APPLICANT:  Wood,William
;  APPLICANT:  Zhang, Zemin
;  TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;  TITLE OF INVENTION:  ACIDS ENCODING THE SAME
;  FILE REFERENCE: P3330R1C160
;  CURRENT APPLICATION NUMBER: US/10/140,470
;  CURRENT FILING DATE:  2002-05-06
;  Prior Application removed - See Palm or File Wrapper
;  NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 362
;   LENGTH: 243
;   TYPE: PRT
;   ORGANISM: Homo Sapien
US-10-140-470-362


  Query Match             100.0%;  Score 243;  DB 9;  Length 243;
  Best Local Similarity   100.0%;  Pred. No. 7.9e-202;
  Matches  243;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MRPLLVLLLLGLAAGSPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MRPLLVLLLLGLAAGSPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPG 60

Qy         61 APGEKGEGGRPGLPGPRGDPGPRGEAGPAGPTGPAGECSVPPRSAFSAKRSESRVPPPSD 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 APGEKGEGGRPGLPGPRGDPGPRGEAGPAGPTGPAGECSVPPRSAFSAKRSESRVPPPSD 120

Qy        121 APLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGESIASFFQ 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 APLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGESIASFFQ 180

Qy        181 FFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWHSSP 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 FFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWHSSP 240

Qy        241 VFA 243
              |||
Db        241 VFA 243


RESULT 11
US-09-796-753-68
```

```
; Sequence 68, Application US/09796753
; Publication No. US20030027998A1
; GENERAL INFORMATION:
;  APPLICANT:  McCarthy, Sean A.
;  TITLE OF INVENTION:  SECRETED PROTEINS AND USES THEREOF
;  FILE REFERENCE:  7853-227-999
;  CURRENT APPLICATION NUMBER: US/09/796,753
;  CURRENT FILING DATE:  2001-03-01
;  PRIOR APPLICATION NUMBER:  09/183,175
;  PRIOR FILING DATE:  1998-10-30
;  PRIOR APPLICATION NUMBER:  09/223,094
;  PRIOR FILING DATE:  1998-12-30
;  PRIOR APPLICATION NUMBER:  09/223,546
;  PRIOR FILING DATE:  1998-12-30
;  PRIOR APPLICATION NUMBER:  09/224,246
;  PRIOR FILING DATE:  1998-12-30
;  PRIOR APPLICATION NUMBER:  09/259,388
;  PRIOR FILING DATE:  1999-02-26
;  PRIOR APPLICATION NUMBER:  60/122,458
;  PRIOR FILING DATE:  1999-03-01
;  PRIOR APPLICATION NUMBER:  09/312,359
;  PRIOR FILING DATE:  1999-05-14
;  PRIOR APPLICATION NUMBER:  09/336,536
;  PRIOR FILING DATE:  1999-06-18
;  PRIOR APPLICATION NUMBER:  09/342,687
;  PRIOR FILING DATE:  1999-06-29
;  PRIOR APPLICATION NUMBER:  09/345,464
;  PRIOR FILING DATE:  1999-06-30
;  PRIOR APPLICATION NUMBER:  09/365,164
;  PRIOR FILING DATE:  1999-07-30
;  PRIOR APPLICATION NUMBER:  09/399,723
;  PRIOR FILING DATE:  1999-09-20
;  PRIOR APPLICATION NUMBER:  09/409,634
;  PRIOR FILING DATE:  1999-09-30
;  PRIOR APPLICATION NUMBER:  09/471,179
;  PRIOR FILING DATE:  1999-12-23
;  PRIOR APPLICATION NUMBER:  09/474,071
;  PRIOR FILING DATE:  1999-12-29
;  PRIOR APPLICATION NUMBER:  09/474,072
;  PRIOR FILING DATE:  1999-12-29
;  PRIOR APPLICATION NUMBER:  09/514,010
;  PRIOR FILING DATE:  2000-02-25
;  PRIOR APPLICATION NUMBER:  09/516,745
;  PRIOR FILING DATE:  2000-03-01
;  PRIOR APPLICATION NUMBER:  09/572,002
;  PRIOR FILING DATE:  2000-05-14
;  PRIOR APPLICATION NUMBER:  09/597,993
;  PRIOR FILING DATE:  2000-06-19
;  PRIOR APPLICATION NUMBER:  09/599,596
;  PRIOR FILING DATE:  2000-06-22
;  PRIOR APPLICATION NUMBER:  09/630,334
;  PRIOR FILING DATE:  2000-07-31
;  PRIOR APPLICATION NUMBER:  09/606,565
;  PRIOR FILING DATE:  2000-06-29
;  PRIOR APPLICATION NUMBER:  09/606,317
;  PRIOR FILING DATE:  2000-06-29
;  PRIOR APPLICATION NUMBER:  09/665,666
;  PRIOR FILING DATE:  2000-09-20
;  PRIOR APPLICATION NUMBER:  09/677,751
;  PRIOR FILING DATE:  2000-09-30
;  NUMBER OF SEQ ID NOS:  162
; SEQ ID NO 68
;   LENGTH: 243
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-796-753-68

  Query Match           100.0%;  Score 243;  DB 9;  Length 243;
  Best Local Similarity  100.0%;  Pred. No. 7.9e-202;
  Matches  243;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MRPLLVLLLLGLAAGSPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db          1 MRPLLVLLLLGLAAGSPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPG 60

Qy         61 APGEKGEGGRPGLPGPRGDPGPRGEAGPAGPTGPAGECSVPPRSAFSAKRSESRVPPPSD 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db         61 APGEKGEGGRPGLPGPRGDPGPRGEAGPAGPTGPAGECSVPPRSAFSAKRSESRVPPPSD 120

Qy        121 APLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGESIASFFQ 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        121 APLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGESIASFFQ 180

Qy        181 FFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWHSSP 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        181 FFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWHSSP 240

Qy        241 VFA 243
              |||

Db        241 VFA 243


RESULT 12
US-10-175-746-362
; Sequence 362, Application US/10175746
; Publication No. US20030027270A1
; GENERAL INFORMATION:
;  APPLICANT: Baker,Kevin P.
;  APPLICANT:  Beresini,Maureen
;  APPLICANT:  DeForge,Laura
;  APPLICANT:  Desnoyers,Luc
;  APPLICANT:  Filvaroff,Ellen
;  APPLICANT:  Gao,Wei-Qiang
;  APPLICANT:  Gerritsen,Mary E.
;  APPLICANT:  Goddard,Audrey
;  APPLICANT:  Godowski,Paul J.
;  APPLICANT:  Gurney,Austin L.
;  APPLICANT:  Sherwood,Steven
;  APPLICANT:  Smith,Victoria
;  APPLICANT:  Stewart,Timothy A.
;  APPLICANT:  Tumas,Daniel
;  APPLICANT:  Watanabe,Colin K
;  APPLICANT:  Wood,William
;  APPLICANT:  Zhang, Zemin
;  TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;  TITLE OF INVENTION:  ACIDS ENCODING THE SAME
;  FILE REFERENCE: P3330R1C353
;  CURRENT APPLICATION NUMBER: US/10/175,746
;  CURRENT FILING DATE:  2002-06-19
;  Prior Application removed - See File Wrapper or Palm
;  NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 362
;   LENGTH: 243
;   TYPE: PRT
;   ORGANISM: Homo Sapien
US-10-175-746-362

  Query Match           100.0%;  Score 243;  DB 9;  Length 243;
  Best Local Similarity  100.0%;  Pred. No. 7.9e-202;
  Matches  243;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MRPLLVLLLLGLAAGSPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db          1 MRPLLVLLLLGLAAGSPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPG 60

Qy         61 APGEKGEGGRPGLPGPRGDPGPRGEAGPAGPTGPAGECSVPPRSAFSAKRSESRVPPPSD 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db         61 APGEKGEGGRPGLPGPRGDPGPRGEAGPAGPTGPAGECSVPPRSAFSAKRSESRVPPPSD 120

Qy        121 APLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGESIASFFQ 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        121 APLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGESIASFFQ 180

Qy        181 FFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWHSSP 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        181 FFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWHSSP 240
```

```
Qy        241 VFA 243
              |||
Db        241 VFA 243


RESULT 13
US-10-176-918-362
; Sequence 362, Application US/10176918
; Publication No. US20030027275A1
; GENERAL INFORMATION:
;  APPLICANT: Baker,Kevin P.
;  APPLICANT:  Beresini,Maureen
;  APPLICANT:  DeForge,Laura
;  APPLICANT:  Desnoyers,Luc
;  APPLICANT:  Filvaroff,Ellen
;  APPLICANT:  Gao,Wei-Qiang
;  APPLICANT:  Gerritsen,Mary E.
;  APPLICANT:  Goddard,Audrey
;  APPLICANT:  Godowski,Paul J.
;  APPLICANT:  Gurney,Austin L.
;  APPLICANT:  Sherwood,Steven
;  APPLICANT:  Smith,Victoria
;  APPLICANT:  Stewart,Timothy A.
;  APPLICANT:  Tumas,Daniel
;  APPLICANT:  Watanabe,Colin K
;  APPLICANT:  Wood,William
;  APPLICANT:  Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION:  ACIDS ENCODING THE SAME
; FILE REFERENCE: P3330R1C382
; CURRENT APPLICATION NUMBER: US/10/176,918
; CURRENT FILING DATE:  2002-06-20
;    Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 362
;   LENGTH: 243
;   TYPE: PRT
;   ORGANISM: Homo Sapien
US-10-176-918-362

  Query Match           100.0%;  Score 243;  DB 9;  Length 243;
  Best Local Similarity  100.0%;  Pred. No. 7.9e-202;
  Matches  243;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MRPLLVLLLLGLAAGSPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPG 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MRPLLVLLLLGLAAGSPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPG 60

Qy         61 APGEKGEGGRPGLPGPRGDPGPRGEAGPAGPTGPAGECSVPPRSAFSAKRSESRVPPPSD 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 APGEKGEGGRPGLPGPRGDPGPRGEAGPAGPTGPAGECSVPPRSAFSAKRSESRVPPPSD 120

Qy        121 APLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGESIASFFQ 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 APLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGESIASFFQ 180

Qy        181 FFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWHSSP 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 FFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWHSSP 240

Qy        241 VFA 243
              |||
Db        241 VFA 243


RESULT 14
US-10-176-921-362
; Sequence 362, Application US/10176921
; Publication No. US20030027276A1
; GENERAL INFORMATION:
;  APPLICANT: Baker,Kevin P.
;  APPLICANT:  Beresini,Maureen
;  APPLICANT:  DeForge,Laura
;  APPLICANT:  Desnoyers,Luc
;  APPLICANT:  Filvaroff,Ellen
;  APPLICANT:  Gao,Wei-Qiang
;  APPLICANT:  Gerritsen,Mary E.
;  APPLICANT:  Goddard,Audrey
;  APPLICANT:  Godowski,Paul J.
;  APPLICANT:  Gurney,Austin L.
;  APPLICANT:  Sherwood,Steven
;  APPLICANT:  Smith,Victoria
;  APPLICANT:  Stewart,Timothy A.
;  APPLICANT:  Tumas,Daniel
;  APPLICANT:  Watanabe,Colin K
;  APPLICANT:  Wood,William
;  APPLICANT:  Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION:  ACIDS ENCODING THE SAME
; FILE REFERENCE: P3330R1C288
; CURRENT APPLICATION NUMBER: US/10/176,921
; CURRENT FILING DATE:  2002-06-20
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 362
;   LENGTH: 243
;   TYPE: PRT
;   ORGANISM: Homo Sapien
US-10-176-921-362

  Query Match           100.0%;  Score 243;  DB 9;  Length 243;
  Best Local Similarity  100.0%;  Pred. No. 7.9e-202;
  Matches  243;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MRPLLVLLLLGLAAGSPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPG 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MRPLLVLLLLGLAAGSPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPG 60

Qy         61 APGEKGEGGRPGLPGPRGDPGPRGEAGPAGPTGPAGECSVPPRSAFSAKRSESRVPPPSD 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 APGEKGEGGRPGLPGPRGDPGPRGEAGPAGPTGPAGECSVPPRSAFSAKRSESRVPPPSD 120

Qy        121 APLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGESIASFFQ 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 APLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGESIASFFQ 180

Qy        181 FFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWHSSP 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 FFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWHSSP 240

Qy        241 VFA 243
              |||
Db        241 VFA 243


RESULT 15
US-10-137-865-362
; Sequence 362, Application US/10137865
; Publication No. US20030032155A1
; GENERAL INFORMATION:
;  APPLICANT: Baker,Kevin P.
;  APPLICANT:  Beresini,Maureen
;  APPLICANT:  DeForge,Laura
;  APPLICANT:  Desnoyers,Luc
;  APPLICANT:  Filvaroff,Ellen
;  APPLICANT:  Gao,Wei-Qiang
;  APPLICANT:  Gerritsen,Mary E.
;  APPLICANT:  Goddard,Audrey
;  APPLICANT:  Godowski,Paul J.
;  APPLICANT:  Gurney,Austin L.
;  APPLICANT:  Sherwood,Steven
;  APPLICANT:  Smith,Victoria
;  APPLICANT:  Stewart,Timothy A.
;  APPLICANT:  Tumas,Daniel
;  APPLICANT:  Watanabe,Colin K
```

; APPLICANT: Wood,William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3330R1C154
; CURRENT APPLICATION NUMBER: US/10/137,865
; CURRENT FILING DATE: 2002-05-03
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 362
; LENGTH: 243
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-137-865-362

Query Match 100.0%; Score 243; DB 9; Length 243;
Best Local Similarity 100.0%; Pred. No. 7.9e-202;
Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MRPLLVLLLLGLAAGSPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MRPLLVLLLLGLAAGSPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPG 60

Qy         61 APGEKGEGGRPGLPGPRGDPGPRGEAGPAGPTGPAGECSVPPRSAFSAKRSESRVPPPSD 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 APGEKGEGGRPGLPGPRGDPGPRGEAGPAGPTGPAGECSVPPRSAFSAKRSESRVPPPSD 120

Qy        121 APLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGESIASFFQ 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 APLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGESIASFFQ 180

Qy        181 FFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWHSSP 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 FFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWHSSP 240

Qy        241 VFA 243
              |||
Db        241 VFA 243
```

Search completed: June 18, 2003, 15:15:19
Job time : 54 secs

GenCore version 5.1.6


OM nucleic - nucleic search, using sw model

Run on: June 22, 2003, 02:43:56 ; Search time 3598 Seconds
(without alignments)
11138.011 Million cell updates/sec

Title: US-09-943-851A-41
Perfect score: 1377
Sequence: 1 gactagttctcttggagtct.....aaaaaaaaaaaaaaaaaaa 1377

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|---|
| 1 | | 1377 | 100.0 | 1377 | 6 | AX464228 | AX464228 Sequence |
| 2 | | 1347.4 | 97.9 | 1370 | 9 | BC029485 | BC029485 Homo sapi |
| 3 | | 1322.2 | 96.0 | 1337 | 9 | AF329841 | AF329841 Homo sapi |
| 4 | | 1322.2 | 96.0 | 1347 | 6 | AR138193 | AR138193 Sequence |
| 5 | | 1064.4 | 77.3 | 1068 | 9 | HSM800923 | AL110261 Homo sapi |
| 6 | c | 909.8 | 66.1 | 191362 | 9 | AP003396 | AP003396 Homo sapi |
| 7 | c | 909.8 | 66.1 | 219574 | 9 | AP002956 | AP002956 Homo sapi |
| 8 | | 896.2 | 65.1 | 182429 | 2 | AP001557 | AP001557 Homo sapi |
| 9 | | 696.2 | 50.6 | 4220 | 10 | AF469650 | AF469650 Mus muscu |
| 10 | | 694 | 50.4 | 173038 | 2 | AP001003 | AP001003 Homo sapi |
| 11 | | 693.6 | 50.4 | 1234 | 10 | BC025174 | BC025174 Mus muscu |
| 12 | | 691.2 | 50.2 | 1271 | 10 | BC023068 | BC023068 Mus muscu |
| 13 | | 612.8 | 44.5 | 198902 | 2 | AP001156 | AP001156 Homo sapi |
| 14 | | 513.4 | 37.3 | 173657 | 2 | AC112557 | AC112557 Rattus no |
| 15 | | 504 | 36.6 | 729 | 6 | AR138195 | AR138195 Sequence |
| 16 | | 500.4 | 36.3 | 249283 | 2 | AC124577 | AC124577 Mus muscu |
| 17 | c | 472.4 | 34.3 | 150653 | 2 | AC107174 | AC107174 Rattus no |
| 18 | | 439.4 | 31.9 | 441 | 9 | HUMZD92F04 | AF086482 Homo sapi |
| 19 | | 187.4 | 13.6 | 150653 | 2 | AC107174 | AC107174 Rattus no |
| 20 | | 161.8 | 11.8 | 2288 | 9 | AK055132 | AK055132 Homo sapi |
| 21 | | 107 | 7.8 | 1886 | 5 | AB067770 | AB067770 Oncorhync |
| 22 | | 100.6 | 7.3 | 4443 | 9 | AK074129 | AK074129 Homo sapi |
| 23 | | 100.6 | 7.3 | 4908 | 6 | AX430953 | AX430953 Sequence |
| 24 | | 100.6 | 7.3 | 128133 | 9 | AL138787 | AL138787 Human DNA |
| 25 | c | 100.6 | 7.3 | 173817 | 2 | AC012141 | AC012141 Homo sapi |
| 26 | | 99.6 | 7.2 | 855 | 6 | AX039974 | AX039974 Sequence |
| 27 | | 97.4 | 7.1 | 1938 | 9 | HUM28CGLE | M60832 Human alpha |
| 28 | c | 92.6 | 6.7 | 152147 | 9 | AC091842 | AC091842 Homo sapi |
| 29 | | 92.6 | 6.7 | 153248 | 9 | AC112191 | AC112191 Homo sapi |
| 30 | | 91.6 | 6.7 | 1161 | 6 | AX039965 | AX039965 Sequence |
| 31 | | 91.6 | 6.7 | 1176 | 9 | AF329836 | AF329836 Homo sapi |
| 32 | | 91.6 | 6.7 | 2404 | 9 | BC011699 | BC011699 Homo sapi |
| 33 | | 91.4 | 6.6 | 1134 | 4 | AF269230 | AF269230 Bos tauru |
| 34 | | 90.6 | 6.6 | 947 | 10 | BC028770 | BC028770 Mus muscu |
| 35 | | 90.6 | 6.6 | 1276 | 6 | AR034252 | AR034252 Sequence |
| 36 | | 90.6 | 6.6 | 1276 | 6 | AX195209 | AX195209 Sequence |
| 37 | | 90.6 | 6.6 | 1276 | 6 | AX358519 | AX358519 Sequence |
| 38 | | 90.6 | 6.6 | 1276 | 10 | MMU37222 | U37222 Mus musculu |
| 39 | c | 90.2 | 6.6 | 194674 | 2 | AC122477 | AC122477 Mus muscu |
| 40 | | 90 | 6.5 | 132884 | 2 | AC128403 | AC128403 Rattus no |
| 41 | | 88.2 | 6.4 | 112554 | 2 | AC129431 | AC129431 Rattus no |
| 42 | | 88 | 6.4 | 1357 | 6 | AX068378 | AX068378 Sequence |
| 43 | | 88 | 6.4 | 1393 | 9 | AF329838 | AF329838 Homo sapi |
| 44 | | 88 | 6.4 | 159766 | 9 | AC104942 | AC104942 Homo sapi |
| 45 | | 86.8 | 6.3 | 178850 | 9 | AL359736 | AL359736 Human DNA |

ALIGNMENTS

RESULT 1
AX464228
LOCUS AX464228 1377 bp DNA linear PAT 16-JUL-2002
DEFINITION Sequence 361 from Patent WO0140466.
ACCESSION AX464228
VERSION AX464228.1 GI:21899130
KEYWORDS .
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Baker,K.P., Beresini,M., Deforge,L., Desnoyers,L., Filvaroff,E., Gao,W.Q., Gerritsen,M.E., Goddard,A., Godowski,P.J., Gurney,A.L., Sherwood,S., Smith,V., Stewart,T.A., Tumas,D., Watanabe,C.K.,

```
  TITLE     Wood,W.L. and Zhang,Z.
            Secreted and transmembrane polypeptides and nucleic acids encoding
            same
  JOURNAL   Patent: WO 0140466-A 361 07-JUN-2001;
            Genentech Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .1377
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT      251 a    422 c    472 g    232 t
ORIGIN

  Query Match          100.0%;  Score 1377;  DB 6;  Length 1377;
  Best Local Similarity  100.0%;  Pred. No. 1.3e-227;
  Matches 1377;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GACTAGTTCTCTTGGAGTCTGGGAGGAGCAAAGCGGAGCCGGCAGGGAGCGAACCAGGAC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GACTAGTTCTCTTGGAGTCTGGGAGGAGCAAAGCGGAGCCGGCAGGGAGCGAACCAGGAC 60

Qy         61 TGGGGTGACGGCAGGGCAGGGGGCGCCTGGCCGGGGAGAAGCGCGGGGGCTGGAGCACCA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TGGGGTGACGGCAGGGCAGGGGGCGCCTGGCCGGGGAGAAGCGCGGGGGCTGGAGCACCA 120

Qy        121 CCAACTGGAGGGTCCGGAGTAGCGAGCGCCCCGAAGGAGGCCATCGGGGAGCCGGGAGGG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CCAACTGGAGGGTCCGGAGTAGCGAGCGCCCCGAAGGAGGCCATCGGGGAGCCGGGAGGG 180

Qy        181 GGGACTGCGAGAGGACCCCGGCGTCCGGGCTCCCGGTGCCAGCGCTATGAGGCCACTCCT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGGACTGCGAGAGGACCCCGGCGTCCGGGCTCCCGGTGCCAGCGCTATGAGGCCACTCCT 240

Qy        241 CGTCCTGCTGCTCCTGGGCCTGGCGGCCGGCTCGCCCCCACTGGACGACAACAAGATCCC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CGTCCTGCTGCTCCTGGGCCTGGCGGCCGGCTCGCCCCCACTGGACGACAACAAGATCCC 300

Qy        301 CAGCCTCTGCCCGGGGCACCCCGGCCTTCCAGGCACGCCGGGCCACCATGGCAGCCAGGG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CAGCCTCTGCCCGGGGCACCCCGGCCTTCCAGGCACGCCGGGCCACCATGGCAGCCAGGG 360

Qy        361 CTTGCCGGGCCGCGATGGCCGCGACGGCCGCGACGGCGCGCCCGGGGCTCCGGGAGAGAA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CTTGCCGGGCCGCGATGGCCGCGACGGCCGCGACGGCGCGCCCGGGGCTCCGGGAGAGAA 420

Qy        421 AGGCGAGGGCGGGAGGCCGGGACTGCCGGGACCTCGAGGGGACCCCGGGCCGCGAGGAGA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AGGCGAGGGCGGGAGGCCGGGACTGCCGGGACCTCGAGGGGACCCCGGGCCGCGAGGAGA 480

Qy        481 GGCGGGACCCGCGGGGCCCACCGGGCCTGCCGGGGAGTGCTCGGTGCCTCCGCGATCCGC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GGCGGGACCCGCGGGGCCCACCGGGCCTGCCGGGGAGTGCTCGGTGCCTCCGCGATCCGC 540

Qy        541 CTTCAGCGCCAAGCGCTCCGAGAGCCGGGTGCCTCCGCCGTCTGACGCACCCTTGCCCTT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CTTCAGCGCCAAGCGCTCCGAGAGCCGGGTGCCTCCGCCGTCTGACGCACCCTTGCCCTT 600

Qy        601 CGACCGCGTGCTGGTGAACGAGCAGGGACATTACGACGCCGTCACCGGCAAGTTCACCTG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CGACCGCGTGCTGGTGAACGAGCAGGGACATTACGACGCCGTCACCGGCAAGTTCACCTG 660

Qy        661 CCAGGTGCCTGGGGTCTACTACTTCGCCGTCCATGCCACCGTCTACCGGGCCAGCCTGCA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CCAGGTGCCTGGGGTCTACTACTTCGCCGTCCATGCCACCGTCTACCGGGCCAGCCTGCA 720

Qy        721 GTTTGATCTGGTGAAGAATGGCGAATCCATTGCCTCTTTCTTCCAGTTTTTCGGGGGGTG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GTTTGATCTGGTGAAGAATGGCGAATCCATTGCCTCTTTCTTCCAGTTTTTCGGGGGGTG 780

Qy        781 GCCCAAGCCAGCCTCGCTCTCGGGGGGGGCCATGGTGAGGCTGGAGCCTGAGGACCAAGT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GCCCAAGCCAGCCTCGCTCTCGGGGGGGGCCATGGTGAGGCTGGAGCCTGAGGACCAAGT 840

Qy        841 GTGGGTGCAGGTGGGTGTGGGTGACTACATTGGCATCTATGCCAGCATCAAGACAGACAG 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GTGGGTGCAGGTGGGTGTGGGTGACTACATTGGCATCTATGCCAGCATCAAGACAGACAG 900

Qy        901 CACCTTCTCCGGATTTCTGGTGTACTCCGACTGGCACAGCTCCCCAGTCTTTGCTTAGTG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 CACCTTCTCCGGATTTCTGGTGTACTCCGACTGGCACAGCTCCCCAGTCTTTGCTTAGTG 960

Qy        961 CCCACTGCAAAGTGAGCTCATGCTCTCACTCCTAGAAGGAGGGTGTGAGGCTGACAACCA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 CCCACTGCAAAGTGAGCTCATGCTCTCACTCCTAGAAGGAGGGTGTGAGGCTGACAACCA 1020

Qy       1021 GGTCATCCAGGAGGGCTGGCCCCCCTGGAATATTGTGAATGACTAGGGAGGTGGGGTAGA 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GGTCATCCAGGAGGGCTGGCCCCCCTGGAATATTGTGAATGACTAGGGAGGTGGGGTAGA 1080

Qy       1081 GCACTCTCCGTCCTGCTGCTGGCAAGGAATGGGAACAGTGGCTGTCTGCGATCAGGTCTG 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 GCACTCTCCGTCCTGCTGCTGGCAAGGAATGGGAACAGTGGCTGTCTGCGATCAGGTCTG 1140

Qy       1141 GCAGCATGGGGCAGTGGCTGGATTTCTGCCCAAGACCAGAGGAGTGTGCTGTGCTGGCAA 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 GCAGCATGGGGCAGTGGCTGGATTTCTGCCCAAGACCAGAGGAGTGTGCTGTGCTGGCAA 1200

Qy       1201 GTGTAAGTCCCCCAGTTGCTCTGGTCCAGGAGCCCACGGTGGGGTGCTCTCTTCCTGGTC 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 GTGTAAGTCCCCCAGTTGCTCTGGTCCAGGAGCCCACGGTGGGGTGCTCTCTTCCTGGTC 1260

Qy       1261 CTCTGCTTCTCTGGATCCTCCCCACCCCCTCCTGCTCCTGGGGCCGGCCCTTTTCTCAGA 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 CTCTGCTTCTCTGGATCCTCCCCACCCCCTCCTGCTCCTGGGGCCGGCCCTTTTCTCAGA 1320

Qy       1321 GATCACTCAATAAACCTAAGAACCCTCATAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1377
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 GATCACTCAATAAACCTAAGAACCCTCATAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1377

RESULT 2
BC029485
LOCUS       BC029485                 1370 bp    mRNA    linear   PRI 16-MAY-2002
DEFINITION  Homo sapiens, C1q and tumor necrosis factor related protein 5,
            clone MGC:32938 IMAGE:5278184, mRNA, complete cds.
ACCESSION   BC029485
VERSION     BC029485.1  GI:20810468
KEYWORDS    MGC.
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1370)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (01-MAY-2002) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
            cDNA Library Preparation: Michael J. Brownstein (NHGRI) &  Shiraki
            Toshiyuki and Piero Carninci (RIKEN)
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Sequencing Group at the Stanford Human Genome
            Center, Stanford University School of Medicine, Stanford, CA  94305
            Web site:       http://www-shgc.stanford.edu
            Contact:  (Dickson, Mark) mcd@paxil.stanford.edu
            Dickson, M., Schmutz, J., Grimwood, J., Rodriquez, A., and Myers;
            R. M.

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAK Plate: 48 Row: h Column: 14
            This clone was selected for full length sequencing because it
```

passed the following selection criteria: matched mRNA gi: 14149711.

```
FEATURES             Location/Qualifiers
     source          1. .1370
                     /organism="Homo sapiens"
                     /db_xref="LocusID:114902"
                     /db_xref="taxon:9606"
                     /clone="MGC:32938 IMAGE:5278184"
                     /tissue_type="Brain, hypothalamus"
                     /clone_lib="NIH_MGC_96"
                     /lab_host="DH10B"
                     /note="Vector: pBluescript"
     CDS             234. .965
                     /codon_start=1
                     /product="C1q and tumor necrosis factor related protein 5"
                     /protein_id="AAH29485.1"
                     /db_xref="GI:20810469"
                     /translation="MRPLLVLLLLGLAAGSPPLDDNKIPSLCPGHPGLPGTPGHHGSR
                     GLPGRDGRDGRDGAPGAPGEKGEGGRPGLPGPRGDPGPRGEAGPAGPTGPAGECSVPP
                     RSAFSAKRSESRVPPPSDAPLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVY
                     RASLQFDLVKNGESIASFFQFFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIY
                     ASIKTDSTFSGFLVYSDWHSSPVFA"
BASE COUNT       239 a    422 c    475 g    234 t
ORIGIN
```

Query Match 97.9%; Score 1347.4; DB 9; Length 1370;
Best Local Similarity 99.6%; Pred. No. 1.6e-222;
Matches 1351; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```
Qy       7 TTCTCTTGGAGTCTGGGAGGAGGAAAGCGGAGCCGGCAGGGAGCGAACCAGGACTGGGGT 66
Db      14 TCCTCTTGGAGTCTGGGAGGAGGAAAGCGGAGCCGGCAGGGAGCGAACCAGGACTGGGGT 73

Qy      67 GACGGCAGGGCAGGGGGCGCCTGGCCGGGGAGAAGCGCGGGGGCTGGAGCACCACCAACT 126
Db      74 GACGGCAGGGCAGGGGGCGCCTGGCCGGGGAGAAGCGCGGGGGCTGGAGCACCACCAACT 133

Qy     127 GGAGGGTCCGGAGTAGCGAGCGCCCCGAAGGAGGCCATCGGGGAGCCGGGAGGGGGGACT 186
Db     134 GGAGGGTCCGGAGTAGCGAGCGCCCCGAAGGAGGCCATCGGGGAGCCGGGAGGGGGGACT 193

Qy     187 GCGAGAGGACCCCGGCGTCCGGGCTCCCGGTGCCAGCGCTATGAGGCCACTCCTCGTCCT 246
Db     194 GCGAGAGGACCCCGGCGTCCGGGCTCCCGGTGCCAGCGCTATGAGGCCACTCCTCGTCCT 253

Qy     247 GCTGCTCCTGGGCCTGGCGGCCGGCTCGCCCCCACTGGACGACAACAAGATCCCCAGCCT 306
Db     254 GCTGCTCCTGGGCCTGGCGGCCGGCTCGCCCCCACTGGACGACAACAAGATCCCCAGCCT 313

Qy     307 CTGCCCGGGGCACCCCGGCCTTCCAGGCACGCCGGGCCACCATGGCAGCCAGGGCTTGCC 366
Db     314 CTGCCCGGGGCACCCCGGCCTTCCAGGCACGCCGGGCCACCATGGCAGCCGGGGCTTGCC 373

Qy     367 GGGCCGCGATGGCCGCGACGGCCGCGACGGCGCGCCCGGGGCTCCGGGAGAGAAAGGCGA 426
Db     374 GGGCCGCGATGGCCGCGACGGCCGCGACGGCGCGCCCGGGGCTCCGGGAGAGAAAGGCGA 433

Qy     427 GGGCGGGAGGCCGGGACTGCCGGGACCTCGAGGGGACCCCGGGCCGCGAGGAGAGGCGGG 486
Db     434 GGGCGGGAGGCCGGGACTGCCGGGACCTCGAGGGGACCCCGGGCCGCGAGGAGAGGCGGG 493

Qy     487 ACCCGCGGGGCCCACCGGGCCTGCCGGGGAGTGCTCGGTGCCTCCGCGATCCGCCTTCAG 546
Db     494 ACCCGCGGGGCCCACCGGGCCTGCCGGGGAGTGCTCGGTGCCTCCGCGATCCGCCTTCAG 553

Qy     547 CGCCAAGCGCTCCGAGAGCCGGGTGCCTCCGCCGTCTGACGCACCCTTGCCCTTCGACCG 606
Db     554 CGCCAAGCGCTCCGAGAGCCGGGTGCCTCCGCCGTCTGACGCACCCTTGCCCTTCGACCG 613

Qy     607 CGTGCTGGTGAACGAGCAGGGACATTACGACGCCGTCACCGGCAAGTTCACCTGCCAGGT 666
Db     614 CGTGCTGGTGAACGAGCAGGGACATTACGACGCCGTCACCGGCAAGTTCACCTGCCAGGT 673

Qy     667 GCCTGGGGTCTACTACTTCGCCGTCCATGCCACCGTCTACCGGGCCAGCCTGCAGTTTGA 726
Db     674 GCCTGGGGTCTACTACTTCGCCGTCCATGCCACCGTCTACCGGGCCAGCCTGCAGTTTGA 733

Qy     727 TCTGGTGAAGAATGGCGAATCCATTGCCTCTTTCTTCCAGTTTTTCGGGGGGTGGCCCAA 786
Db     734 TCTGGTGAAGAATGGCGAATCCATTGCCTCTTTCTTCCAGTTTTTCGGGGGGTGGCCCAA 793

Qy     787 GCCAGCCTCGCTCTCGGGGGGGGCCATGGTGAGGCTGGAGCCTGAGGACCAAGTGTGGGT 846
Db     794 GCCAGCCTCGCTCTCGGGGGGGGCCATGGTGAGGCTGGAGCCTGAGGACCAAGTGTGGGT 853

Qy     847 GCAGGTGGGTGTGGGTGACTACATTGGCATCTATGCCAGCATCAAGACAGACAGCACCTT 906
Db     854 GCAGGTGGGTGTGGGTGACTACATTGGCATCTATGCCAGCATCAAGACAGACAGCACCTT 913

Qy     907 CTCCGGATTTCTGGTGTACTCCGACTGGCACAGCTCCCCAGTCTTTGCTTAGTGCCCACT 966
Db     914 CTCCGGATTTCTGGTGTACTCCGACTGGCACAGCTCCCCAGTCTTTGCTTAGTGCCCACT 973

Qy     967 GCAAAGTGAGCTCATGCTCTCACTCCTAGAAGGAGGGTGTGAGGCTGACAACCAGGTCAT 1026
Db     974 GCAAAGTGAGCTCATGCTCTCACTCCTAGAAGGAGGGTGTGAGGCTGACAACCTGGTCAT 1033

Qy    1027 CCAGGAGGGCTGGCCCCCCTGGAATATTGTGAATGACTAGGGAGGTGGGGTAGAGCACTC 1086
Db    1034 CCAGGAGGGCTGGCCCTCCTGGAATATTGTGAATGACTAGGGAGGTGGGGTAGAGCACTC 1093

Qy    1087 TCCGTCCTGCTGCTGGCAAGGAATGGGAACAGTGGCTGTCTGCGATCAGGTCTGGCAGCA 1146
Db    1094 TCCGTCCTGCTGCTGGCAAGGAATGGGAACAGTGGCTGTCTGCGATCAGGTCTGGCAGCA 1153

Qy    1147 TGGGGCAGTGGCTGGATTTCTGCCCAAGACCAGAGGAGTGTGCTGTGCTGGCAAGTGTAA 1206
Db    1154 TGGGGCAGTGGCTGGATTTCTGCCCAAGACCAGAGGAGTGTGCTGTGCTGGCAAGTGTAA 1213

Qy    1207 GTCCCCCAGTTGCTCTGGTCCAGGAGCCCACGGTGGGGTGCTCTCTTCCTGGTCCTCTGC 1266
Db    1214 GTCCCACAGTTGCTCTGGTCCAGGAGCCCACGGTGGGGTGCTCTCTTCCTGGTCCTCTGC 1273

Qy    1267 TTCTCTGGATCCTCCCCACCCCCTCCTGCTCCTGGGGCCGGCCCTTTTCTCAGAGATCAC 1326
Db    1274 TTCTCTGGATCCTCCCCACCCCCTCCTGCTCCTGGGGCCGGCCCTTTTCTCAGAGATCAC 1333

Qy    1327 TCAATAAACCTAAGAACCCTCATAAAAAAAAAAAAAA 1363
Db    1334 TCAATAAACCTAAGAACCCTCAAAAAAAAAAAAAAA 1370
```

```
RESULT 3
AF329841
LOCUS       AF329841               1337 bp    mRNA    linear   PRI 12-MAR-2001
DEFINITION  Homo sapiens complement-clq tumor necrosis factor-related protein
            (CTRP5) mRNA, complete cds.
ACCESSION   AF329841
VERSION     AF329841.1  GI:13274527
KEYWORDS    .
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1337)
  AUTHORS   Sheppard,P.O. and Humes,J.M.
  TITLE     Homo sapiens complement-clq tumor necrosis factor-related protein
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1337)
  AUTHORS   Sheppard,P.O. and Humes,J.M.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-DEC-2000) Bioinformatics, ZymoGenetics, Inc., 1201
            Eastlake Ave. East, Seattle, WA 98102, USA
FEATURES             Location/Qualifiers
     source          1. .1337
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
```

```
                     /chromosome="11"
                     /map="11q23.3"
     gene            1. .1337
                     /gene="CTRP5"
                     /note="ZSIG39"
     CDS             198. .929
                     /gene="CTRP5"
                     /codon_start=1
                     /product="complement-c1q tumor necrosis factor-related
                     protein"
                     /protein_id="AAK17965.1"
                     /db_xref="GI:13274528"
                     /translation="MRPLLVLLLLGLAAGSPPLDDNKIPSLCPGHPGLPGTPGHHGSQ
                     GLPGRDGRDGRDGAPGAPGEKGEGGRPGLPGPRGDPGPRGEAGPAGPTGPAGECSVPP
                     RSAFSAKRSESRVPPPSDAPLPFDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVY
                     RASLQFDLVKNGESIASFFQFFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIY
                     ASIKTDSTFSGFLVYSDWHSSPVFA"
BASE COUNT       235 a    417 c    459 g    226 t
ORIGIN

  Query Match           96.0%;  Score 1322.2;  DB 9;  Length 1337;
  Best Local Similarity  99.8%;  Pred. No. 3.5e-218;
  Matches 1324;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy         40 CGGCAGGGAGCGAACCAGGACTGGGGTGACGGCAGGGCAGGGGGCGCCTGGCCGGGGAGA 99
               ||  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         11 CGAGAGGGAGCGAACCAGGACTGGGGTGACGGCAGGGCAGGGGGCGCCTGGCCGGGGAGA 70

Qy        100 AGCGCGGGGGCTGGAGCACCACCAACTGGAGGGTCCGGAGTAGCGAGCGCCCCGAAGGAG 159
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         71 AGCGCGGGGGCTGGAGCACCACCAACTGGAGGGTCCGGAGTAGCGAGCGCCCCGAAGGAG 130

Qy        160 GCCATCGGGGAGCCGGGAGGGGGGACTGCGAGAGGACCCCGGCGTCCGGGCTCCCGGTGC 219
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        131 GCCATCGGGGAGCCGGGAGGGGGGACTGCGAGAGGACCCCGGCGTCCGGGCTCCCGGTGC 190

Qy        220 CAGCGCTATGAGGCCACTCCTCGTCCTGCTGCTCCTGGGCCTGGCGGCCGGCTCGCCCCC 279
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        191 CAGCGCTATGAGGCCACTCCTCGTCCTGCTGCTCCTGGGCCTGGCGGCCGGCTCGCCCCC 250

Qy        280 ACTGGACGACAACAAGATCCCCAGCCTCTGCCCGGGGCACCCCGGCCTTCCAGGCACGCC 339
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        251 ACTGGACGACAACAAGATCCCCAGCCTCTGCCCGGGGCACCCCGGCCTTCCAGGCACGCC 310

Qy        340 GGGCCACCATGGCAGCCAGGGCTTGCCGGGCCGCGATGGCCGCGACGGCCGCGACGGCGC 399
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        311 GGGCCACCATGGCAGCCAGGGCTTGCCGGGCCGCGATGGCCGCGACGGCCGCGACGGCGC 370

Qy        400 GCCCGGGGCTCCGGGAGAGAAAGGCGAGGGCGGGAGGCCGGGACTGCCGGGACCTCGAGG 459
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        371 GCCCGGGGCTCCGGGAGAGAAAGGCGAGGGCGGGAGGCCGGGACTGCCGGGACCTCGAGG 430

Qy        460 GGACCCCGGGCCGCGAGGAGAGGCGGGACCCGCGGGGCCCACCGGGCCTGCCGGGGAGTG 519
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        431 GGACCCCGGGCCGCGAGGAGAGGCGGGACCCGCGGGGCCCACCGGGCCTGCCGGGGAGTG 490

Qy        520 CTCGGTGCCTCCGCGATCCGCCTTCAGCGCCAAGCGCTCCGAGAGCCGGGTGCCTCCGCC 579
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        491 CTCGGTGCCTCCGCGATCCGCCTTCAGCGCCAAGCGCTCCGAGAGCCGGGTGCCTCCGCC 550

Qy        580 GTCTGACGCACCCTTGCCCTTCGACCGCGTGCTGGTGAACGAGCAGGGACATTACGACGC 639
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        551 GTCTGACGCACCCTTGCCCTTCGACCGCGTGCTGGTGAACGAGCAGGGACATTACGACGC 610

Qy        640 CGTCACCGGCAAGTTCACCTGCCAGGTGCCTGGGGTCTACTACTTCGCCGTCCATGCCAC 699
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        611 CGTCACCGGCAAGTTCACCTGCCAGGTGCCTGGGGTCTACTACTTCGCCGTCCATGCCAC 670

Qy        700 CGTCTACCGGGCCAGCCTGCAGTTTGATCTGGTGAAGAATGGCGAATCCATTGCCTCTTT 759
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        671 CGTCTACCGGGCCAGCCTGCAGTTTGATCTGGTGAAGAATGGCGAATCCATTGCCTCTTT 730

Qy        760 CTTCCAGTTTTTCGGGGGGTGGCCCAAGCCAGCCTCGCTCTCGGGGGGGGCCATGGTGAG 819
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        731 CTTCCAGTTTTTCGGGGGGTGGCCCAAGCCAGCCTCGCTCTCGGGGGGGGCCATGGTGAG 790

Qy        820 GCTGGAGCCTGAGGACCAAGTGTGGGTGCAGGTGGGTGTGGGTGACTACATTGGCATCTA 879
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        791 GCTGGAGCCTGAGGACCAAGTGTGGGTGCAGGTGGGTGTGGGTGACTACATTGGCATCTA 850

Qy        880 TGCCAGCATCAAGACAGACAGCACCTTCTCCGGATTTCTGGTGTACTCCGACTGGCACAG 939
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        851 TGCCAGCATCAAGACAGACAGCACCTTCTCCGGATTTCTGGTGTACTCCGACTGGCACAG 910

Qy        940 CTCCCCAGTCTTTGCTTAGTGCCCACTGCAAAGTGAGCTCATGCTCTCACTCCTAGAAGG 999
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        911 CTCCCCAGTCTTTGCTTAGTGCCCACTGCAAAGTGAGCTCATGCTCTCACTCCTAGAAGG 970

Qy       1000 AGGGTGTGAGGCTGACAACCAGGTCATCCAGGAGGGCTGGCCCCCCTGGAATATTGTGAA 1059
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        971 AGGGTGTGAGGCTGACAACCAGGTCATCCAGGAGGGCTGGCCCCCCTGGAATATTGTGAA 1030

Qy       1060 TGACTAGGGAGGTGGGGTAGAGCACTCTCCGTCCTGCTGCTGGCAAGGAATGGGAACAGT 1119
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1031 TGACTAGGGAGGTGGGGTAGAGCACTCTCCGTCCTGCTGCTGGCAAGGAATGGGAACAGT 1090

Qy       1120 GGCTGTCTGCGATCAGGTCTGGCAGCATGGGGCAGTGGCTGGATTTCTGCCCAAGACCAG 1179
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1091 GGCTGTCTGCGATCAGGTCTGGCAGCATGGGGCAGTGGCTGGATTTCTGCCCAAGACCAG 1150

Qy       1180 AGGAGTGTGCTGTGCTGGCAAGTGTAAGTCCCCCAGTTGCTCTGGTCCAGGAGCCCACGG 1239
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1151 AGGAGTGTGCTGTGCTGGCAAGTGTAAGTCCCCCAGTTGCTCTGGTCCAGGAGCCCACGG 1210

Qy       1240 TGGGGTGCTCTCTTCCTGGTCCTCTGCTTCTCTGGATCCTCCCCACCCCCTCCTGCTCCT 1299
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1211 TGGGGTGCTCTCTTCCTGGTCCTCTGCTTCTCTGGATCCTCCCCACCCCCTCCTGCTCCT 1270

Qy       1300 GGGGCCGGCCCTTTTCTCAGAGATCACTCAATAAACCTAAGAACCCTCATAAAAAAAAAA 1359
              ||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db       1271 GGGGCCGGCCCTTTTCTCAGAGATCACTCAATAAACCTAAGAACCCTCAAAAAAAAAAAA 1330

Qy       1360 AAAAAAA 1366
              |||||||
Db       1331 AAAAAAA 1337


RESULT 4
AR138193
LOCUS       AR138193                1347 bp    DNA     linear   PAT 16-JUN-2001
DEFINITION  Sequence 1 from patent US 6197930.
ACCESSION   AR138193
VERSION     AR138193.1  GI:14479702
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1347)
  AUTHORS   Sheppard,P.O. and Humes,J.M.
  TITLE     Adipocyte-specific protein homologs
  JOURNAL   Patent: US 6197930-A 1 06-MAR-2001;
FEATURES             Location/Qualifiers
     source          1. .1347
                     /organism="unknown"
BASE COUNT       235 a    421 c    465 g    226 t
ORIGIN

  Query Match           96.0%;  Score 1322.2;  DB 6;  Length 1347;
  Best Local Similarity  99.8%;  Pred. No. 3.5e-218;
  Matches 1324;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy         40 CGGCAGGGAGCGAACCAGGACTGGGGTGACGGCAGGGCAGGGGGCGCCTGGCCGGGGAGA 99
               ||  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         11 CGAGAGGGAGCGAACCAGGACTGGGGTGACGGCAGGGCAGGGGGCGCCTGGCCGGGGAGA 70

Qy        100 AGCGCGGGGGCTGGAGCACCACCAACTGGAGGGTCCGGAGTAGCGAGCGCCCCGAAGGAG 159
```

```
Db        71 AGCGCGGGGGCTGGAGCACCACCAACTGGAGGGTCCGGAGTAGCGAGCGCCCCGAAGGAG 130
Qy       160 GCCATCGGGGAGCCGGGAGGGGGGACTGCGAGAGGACCCCGGCGTCCGGGCTCCCGGTGC 219
Db       131 GCCATCGGGGAGCCGGGAGGGGGGACTGCGAGAGGACCCCGGCGTCCGGGCTCCCGGTGC 190
Qy       220 CAGCGCTATGAGGCCACTCCTCGTCCTGCTGCTCCTGGGCCTGGCGGCCGGCTCGCCCCC 279
Db       191 CAGCGCTATGAGGCCACTCCTCGTCCTGCTGCTCCTGGGCCTGGCGGCCGGCTCGCCCCC 250
Qy       280 ACTGGACGACAACAAGATCCCCAGCCTCTGCCCGGGGCACCCCGGCCTTCCAGGCACGCC 339
Db       251 ACTGGACGACAACAAGATCCCCAGCCTCTGCCCGGGGCACCCCGGCCTTCCAGGCACGCC 310
Qy       340 GGGCCACCATGGCAGCCAGGGCTTGCCGGGCCGCGATGGCCGCGACGGCCGCGACGGCGC 399
Db       311 GGGCCACCATGGCAGCCAGGGCTTGCCGGGCCGCGATGGCCGCGACGGCCGCGACGGCGC 370
Qy       400 GCCCGGGGCTCCGGGAGAGAAAGGCGAGGGCGGGAGGCCGGGACTGCCGGGACCTCGAGG 459
Db       371 GCCCGGGGCTCCGGGAGAGAAAGGCGAGGGCGGGAGGCCGGGACTGCCGGGACCTCGAGG 430
Qy       460 GGACCCCGGGCCGCGAGGAGAGGCGGGACCCGCGGGGCCCACCGGGCCTGCCGGGGAGTG 519
Db       431 GGACCCCGGGCCGCGAGGAGAGGCGGGACCCGCGGGGCCCACCGGGCCTGCCGGGGAGTG 490
Qy       520 CTCGGTGCCTCCGCGATCCGCCTTCAGCGCCAAGCGCTCCGAGAGCCGGGTGCCTCCGCC 579
Db       491 CTCGGTGCCTCCGCGATCCGCCTTCAGCGCCAAGCGCTCCGAGAGCCGGGTGCCTCCGCC 550
Qy       580 GTCTGACGCACCCTTGCCCTTCGACCGCGTGCTGGTGAACGAGCAGGGACATTACGACGC 639
Db       551 GTCTGACGCACCCTTGCCCTTCGACCGCGTGCTGGTGAACGAGCAGGGACATTACGACGC 610
Qy       640 CGTCACCGGCAAGTTCACCTGCCAGGTGCCTGGGGTCTACTACTTCGCCGTCCATGCCAC 699
Db       611 CGTCACCGGCAAGTTCACCTGCCAGGTGCCTGGGGTCTACTACTTCGCCGTCCATGCCAC 670
Qy       700 CGTCTACCGGGCCAGCCTGCAGTTTGATCTGGTGAAGAATGGCGAATCCATTGCCTCTTT 759
Db       671 CGTCTACCGGGCCAGCCTGCAGTTTGATCTGGTGAAGAATGGCGAATCCATTGCCTCTTT 730
Qy       760 CTTCCAGTTTTTCGGGGGGTGGCCCAAGCCAGCCTCGCTCTCGGGGGGGGCCATGGTGAG 819
Db       731 CTTCCAGTTTTTCGGGGGGTGGCCCAAGCCAGCCTCGCTCTCGGGGGGGGCCATGGTGAG 790
Qy       820 GCTGGAGCCTGAGGACCAAGTGTGGGTGCAGGTGGGTGTGGGTGACTACATTGGCATCTA 879
Db       791 GCTGGAGCCTGAGGACCAAGTGTGGGTGCAGGTGGGTGTGGGTGACTACATTGGCATCTA 850
Qy       880 TGCCAGCATCAAGACAGACAGCACCTTCTCCGGATTTCTGGTGTACTCCGACTGGCACAG 939
Db       851 TGCCAGCATCAAGACAGACAGCACCTTCTCCGGATTTCTGGTGTACTCCGACTGGCACAG 910
Qy       940 CTCCCCAGTCTTTGCTTAGTGCCCACTGCAAAGTGAGCTCATGCTCTCACTCCTAGAAGG 999
Db       911 CTCCCCAGTCTTTGCTTAGTGCCCACTGCAAAGTGAGCTCATGCTCTCACTCCTAGAAGG 970
Qy      1000 AGGGTGTGAGGCTGACAACCAGGTCATCCAGGAGGGCTGGCCCCCCTGGAATATTGTGAA 1059
Db       971 AGGGTGTGAGGCTGACAACCAGGTCATCCAGGAGGGCTGGCCCCCCTGGAATATTGTGAA 1030
Qy      1060 TGACTAGGGAGGTGGGGTAGAGCACTCTCCGTCCTGCTGCTGGCAAGGAATGGGAACAGT 1119
Db      1031 TGACTAGGGAGGTGGGGTAGAGCACTCTCCGTCCTGCTGCTGGCAAGGAATGGGAACAGT 1090
Qy      1120 GGCTGTCTGCGATCAGGTCTGGCAGCATGGGGCAGTGGCTGGATTTCTGCCCAAGACCAG 1179
Db      1091 GGCTGTCTGCGATCAGGTCTGGCAGCATGGGGCAGTGGCTGGATTTCTGCCCAAGACCAG 1150
Qy      1180 AGGAGTGTGCTGTGCTGGCAAGTGTAAGTCCCCCAGTTGCTCTGGTCCAGGAGCCCACGG 1239
Db      1151 AGGAGTGTGCTGTGCTGGCAAGTGTAAGTCCCCCAGTTGCTCTGGTCCAGGAGCCCACGG 1210
Qy      1240 TGGGGTGCTCTCTTCCTGGTCCTCTGCTTCTCTGGATCCTCCCCCACCCCCTCCTGCTCCT 1299
Db      1211 TGGGGTGCTCTCTTCCTGGTCCTCTGCTTCTCTGGATCCTCCCCCACCCCCTCCTGCTCCT 1270
Qy      1300 GGGGCCGGCCCTTTTCTCAGAGATCACTCAATAAACCTAAGAACCCTCATAAAAAAAAAAA 1359
Db      1271 GGGGCCGGCCCTTTTCTCAGAGATCACTCAATAAACCTAAGAACCCTCAAAAAAAAAAAA 1330
Qy      1360 AAAAAAA 1366
Db      1331 AAAAAAA 1337
```

```
RESULT 5
HSM800923
LOCUS       HSM800923              1068 bp    mRNA    linear   PRI 18-FEB-2000
DEFINITION  Homo sapiens mRNA; cDNA DKFZp586B0621 (from clone DKFZp586B0621).
ACCESSION   AL110261
VERSION     AL110261.1  GI:5817222
KEYWORDS    .
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1068)
  AUTHORS   Ottenwaelder,B., Obermaier,B., Mewes,H.W., Gassenhuber,J. and
            Wiemann,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-AUG-1999) MIPS, Am Klopferspitz 18a, D-82152
            Martinsried, GERMANY
COMMENT     Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
            Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
            sequenced by MediGenomix (Martinsried/Germany) within the cDNA
            sequencing consortium of the German Genome Project. This clone
            (DKFZp586B0621) is available at the RZPD in Berlin. Please contact
            the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
            Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further
            information about the clone and the sequencing project is available
            at http://www.mips.biochem.mpg.de/proj/cDNA/.
FEATURES             Location/Qualifiers
     source          1. .1068
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="DKFZp586B0621"
                     /tissue_type="uterus"
                     /clone_lib="586 (synonym: hutel). Vector pSport1; host
                     DH10B; sites NotI + SalI/MluI"
                     /dev_stage="adult"
     gene            1. .660
                     /gene="DKFZp586B0621"
     CDS             <1. .660
                     /gene="DKFZp586B0621"
                     /note="similarity to complement subcomponent C1q chain C"
                     /codon_start=1
                     /product="hypothetical protein"
                     /protein_id="CAB53702.1"
                     /db_xref="GI:5817223"
                     /translation="PSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPGAPGEKGEG
                     GRPGLPGPRGDPGPRGEAGPAGPTGPAGECSVPPRSAFSAKRSESRVPPPSDAPLPFD
                     RVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGESIASFFQFFGG
                     WPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIYASIKTDSTFSGFLVYSDWHSSPVF
                     A"
     polyA_signal    1031. .1036
     polyA_site      1049
BASE COUNT      187 a    334 c    349 g    198 t
ORIGIN
```

```
Query Match             77.3%;  Score 1064.4;  DB 9;  Length 1068;
Best Local Similarity   99.9%;  Pred. No. 9.2e-174;
Matches 1065;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;
```

```
Qy        299 CCCAGCCTCTGCCCGGGGCACCCCGGCCTTCCAGGCACGCCGGGCCACCATGGCAGCCAG 358
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CCCAGCCTCTGCCCGGGGCACCCCGGCCTTCCAGGCACGCCGGGCCACCATGGCAGCCAG  60

Qy        359 GGCTTGCCGGGCCGCGATGGCCGCGACGGCCGCGACGGCGCGCCCGGGGCTCCGGGAGAG 418
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GGCTTGCCGGGCCGCGATGGCCGCGACGGCCGCGACGGCGCGCCCGGGGCTCCGGGAGAG 120

Qy        419 AAAGGCGAGGGCGGGAGGCCGGGACTGCCGGGACCTCGAGGGGACCCCGGGCCGCGAGGA 478
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AAAGGCGAGGGCGGGAGGCCGGGACTGCCGGGACCTCGAGGGGACCCCGGGCCGCGAGGA 180

Qy        479 GAGGCGGGACCCGCGGGGCCCACCGGGCCTGCCGGGGAGTGCTCGGTGCCTCCGCGATCC 538
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GAGGCGGGACCCGCGGGGCCCACCGGGCCTGCCGGGGAGTGCTCGGTGCCTCCGCGATCC 240

Qy        539 GCCTTCAGCGCCAAGCGCTCCGAGAGCCGGGTGCCTCCGCCGTCTGACGCACCCTTGCCC 598
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GCCTTCAGCGCCAAGCGCTCCGAGAGCCGGGTGCCTCCGCCGTCTGACGCACCCTTGCCC 300

Qy        599 TTCGACCGCGTGCTGGTGAACGAGCAGGGACATTACGACGCCGTCACCGGCAAGTTCACC 658
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TTCGACCGCGTGCTGGTGAACGAGCAGGGACATTACGACGCCGTCACCGGCAAGTTCACC 360

Qy        659 TGCCAGGTGCCTGGGGTCTACTACTTCGCCGTCCATGCCACCGTCTACCGGGCCAGCCTG 718
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TGCCAGGTGCCTGGGGTCTACTACTTCGCCGTCCATGCCACCGTCTACCGGGCCAGCCTG 420

Qy        719 CAGTTTGATCTGGTGAAGAATGGCGAATCCATTGCCTCTTTCTTCCAGTTTTTCGGGGGG 778
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CAGTTTGATCTGGTGAAGAATGGCGAATCCATTGCCTCTTTCTTCCAGTTTTTCGGGGGG 480

Qy        779 TGGCCCAAGCCAGCCTCGCTCTCGGGGGGGGCCATGGTGAGGCTGGAGCCTGAGGACCAA 838
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TGGCCCAAGCCAGCCTCGCTCTCGGGGGGGGCCATGGTGAGGCTGGAGCCTGAGGACCAA 540

Qy        839 GTGTGGGTGCAGGTGGGTGTGGGTGACTACATTGGCATCTATGCCAGCATCAAGACAGAC 898
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GTGTGGGTGCAGGTGGGTGTGGGTGACTACATTGGCATCTATGCCAGCATCAAGACAGAC 600

Qy        899 AGCACCTTCTCCGGATTTCTGGTGTACTCCGACTGGCACAGCTCCCCAGTCTTTGCTTAG 958
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AGCACCTTCTCCGGATTTCTGGTGTACTCCGACTGGCACAGCTCCCCAGTCTTTGCTTAG 660

Qy        959 TGCCCACTGCAAAGTGAGCTCATGCTCTCACTCCTAGAAGGAGGGTGTGAGGCTGACAAC 1018
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TGCCCACTGCAAAGTGAGCTCATGCTCTCACTCCTAGAAGGAGGGTGTGAGGCTGACAAC 720

Qy       1019 CAGGTCATCCAGGAGGGCTGGCCCCCCTGGAATATTGTGAATGACTAGGGAGGTGGGGTA 1078
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CAGGTCATCCAGGAGGGCTGGCCCCCCTGGAATATTGTGAATGACTAGGGAGGTGGGGTA 780

Qy       1079 GAGCACTCTCCGTCCTGCTGCTGGCAAGGAATGGGAACAGTGGCTGTCTGCGATCAGGTC 1138
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GAGCACTCTCCGTCCTGCTGCTGGCAAGGAATGGGAACAGTGGCTGTCTGCGATCAGGTC 840

Qy       1139 TGGCAGCATGGGGCAGTGGCTGGATTTCTGCCCAAGACCAGAGGAGTGTGCTGTGCTGGC 1198
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 TGGCAGCATGGGGCAGTGGCTGGATTTCTGCCCAAGACCAGAGGAGTGTGCTGTGCTGGC 900

Qy       1199 AAGTGTAAGTCCCCCAGTTGCTCTGGTCCAGGAGCCCACGGTGGGGTGCTCTCTTCCTGG 1258
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 AAGTGTAAGTCCCCCAGTTGCTCTGGTCCAGGAGCCCACGGTGGGGTGCTCTCTTCCTGG 960

Qy       1259 TCCTCTGCTTCTCTGGATCCTCCCCACCCCCTCCTGCTCCTGGGGCCGGCCCTTTTCTCA 1318
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 TCCTCTGCTTCTCTGGATCCTCCCCACCCCCTCCTGCTCCTGGGGCCGGCCCTTTTCTCA 1020

Qy       1319 GAGATCACTCAATAAACCTAAGAACCCTCATAAAAAAAAAAAAAAAA 1364
              |||||||||||||||||||||||||||| ||||||||||||||||
Db       1021 GAGATCACTCAATAAACCTAAGAACCCTCAAAAAAAAAAAAAAAAA 1066
```

```
RESULT 6
AP003396/c
LOCUS       AP003396              191362 bp    DNA     linear   PRI 25-APR-2002
DEFINITION  Homo sapiens genomic DNA, chromosome 11q clone:RP11-334E6, complete
            sequences.
ACCESSION   AP003396
VERSION     AP003396.1  GI:20302607
KEYWORDS    HTG.
SOURCE      Homo sapiens DNA, clone:RP11-334E6.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
            Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
  TITLE     Homo sapiens genomic DNA
  JOURNAL   Published Only in Database (2002)
REFERENCE   2  (bases 1 to 191362)
  AUTHORS   Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
            Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-MAR-2001) Masahira Hattori, The Institute of Physical
            and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
            1-7-22 Suehiro-chou,Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            (E-mail:hattori@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
            Tel:81-45-503-9111, Fax:81-45-503-9170)
COMMENT     This work was done in  collaboration with Arai, Y., Kubo, T. and
            Ohki, M.
            National Cancer Center Research Institute
            Cancer Genomic Division
            5-1-1, Tsukiji, Chuo-ku, Tokyo, JAPAN
            zip: 104-0045
            phone: 81-3-3542-2511 ex 4752, fax: 81-3-3542-0688 e-mail:
            yarai@ncc.go.jp.
FEATURES             Location/Qualifiers
     source          1. .191362
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="11"
                     /map="11q"
                     /clone="RP11-334E6"
BASE COUNT    46535 a  45977 c  48722 g  50128 t
ORIGIN
```

Query Match 66.1%; Score 909.8; DB 9; Length 191362;
Best Local Similarity 99.8%; Pred. No. 2.6e-147;
Matches 911; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy        438 CGGGACTGCCGGGACCTCGAGGGGACCCCGGGCCGCGAGGAGAGGCGGGACCCGCGGGGC 497
              | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6384 CAGGACTGCCGGGACCTCGAGGGGACCCCGGGCCGCGAGGAGAGGCGGGACCCGCGGGGC 6325

Qy        498 CCACCGGGCCTGCCGGGGAGTGCTCGGTGCCTCCGCGATCCGCCTTCAGCGCCAAGCGCT 557
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6324 CCACCGGGCCTGCCGGGGAGTGCTCGGTGCCTCCGCGATCCGCCTTCAGCGCCAAGCGCT 6265

Qy        558 CCGAGAGCCGGGTGCCTCCGCCGTCTGACGCACCCTTGCCCTTCGACCGCGTGCTGGTGA 617
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6264 CCGAGAGCCGGGTGCCTCCGCCGTCTGACGCACCCTTGCCCTTCGACCGCGTGCTGGTGA 6205

Qy        618 ACGAGCAGGGACATTACGACGCCGTCACCGGCAAGTTCACCTGCCAGGTGCCTGGGGTCT 677
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6204 ACGAGCAGGGACATTACGACGCCGTCACCGGCAAGTTCACCTGCCAGGTGCCTGGGGTCT 6145

Qy        678 ACTACTTCGCCGTCCATGCCACCGTCTACCGGGCCAGCCTGCAGTTTGATCTGGTGAAGA 737
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6144 ACTACTTCGCCGTCCATGCCACCGTCTACCGGGCCAGCCTGCAGTTTGATCTGGTGAAGA 6085

Qy        738 ATGGCGAATCCATTGCCTCTTTCTTCCAGTTTTTCGGGGGGTGGCCCAAGCCAGCCTCGC 797
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6084 ATGGCGAATCCATTGCCTCTTTCTTCCAGTTTTTCGGGGGGTGGCCCAAGCCAGCCTCGC 6025

Qy        798 TCTCGGGGGGGGCCATGGTGAGGCTGGAGCCTGAGGACCAAGTGTGGGTGCAGGTGGGTG 857
```

```
Db       6024 TCTCGGGGGGGGCCATGGTGAGGCTGGAGCCTGAGGACCAAGTGTGGGTGCAGGTGGGTG 5965

Qy        858 TGGGTGACTACATTGGCATCTATGCCAGCATCAAGACAGACAGCACCTTCTCCGGATTTC 917

Db       5964 TGGGTGACTACATTGGCATCTATGCCAGCATCAAGACAGACAGCACCTTCTCCGGATTTC 5905

Qy        918 TGGTGTACTCCGACTGGCACAGCTCCCCAGTCTTTGCTTAGTGCCCACTGCAAAGTGAGC 977

Db       5904 TGGTGTACTCCGACTGGCACAGCTCCCCAGTCTTTGCTTAGTGCCCACTGCAAAGTGAGC 5845

Qy        978 TCATGCTCTCACTCCTAGAAGGAGGGTGTGAGGCTGACAACCAGGTCATCCAGGAGGGCT 1037

Db       5844 TCATGCTCTCACTCCTAGAAGGAGGGTGTGAGGCTGACAACCAGGTCATCCAGGAGGGCT 5785

Qy       1038 GGCCCCCCTGGAATATTGTGAATGACTAGGGAGGTGGGGTAGAGCACTCTCCGTCCTGCT 1097

Db       5784 GGCCCCCCTGGAATATTGTGAATGACTAGGGAGGTGGGGTAGAGCACTCTCCGTCCTGCT 5725

Qy       1098 GCTGGCAAGGAATGGGAACAGTGGCTGTCTGCGATCAGGTCTGGCAGCATGGGGCAGTGG 1157

Db       5724 GCTGGCAAGGAATGGGAACAGTGGCTGTCTGCGATCAGGTCTGGCAGCATGGGGCAGTGG 5665

Qy       1158 CTGGATTTCTGCCCAAGACCAGAGGAGTGTGCTGTGCTGGCAAGTGTAAGTCCCCCAGTT 1217

Db       5664 CTGGATTTCTGCCCAAGACCAGAGGAGTGTGCTGTGCTGGCAAGTGTAAGTCCCCCAGTT 5605

Qy       1218 GCTCTGGTCCAGGAGCCCACGGTGGGGTGCTCTCTTCCTGGTCCTCTGCTTCTCTGGATC 1277

Db       5604 GCTCTGGTCCAGGAGCCCACGGTGGGGTGCTCTCTTCCTGGTCCTCTGCTTCTCTGGATC 5545

Qy       1278 CTCCCCACCCCCTCCTGCTCCTGGGGCCGGCCCTTTTCTCAGAGATCACTCAATAAACCT 1337

Db       5544 CTCCCCACCCCCTCCTGCTCCTGGGGCCGGCCCTTTTCTCAGAGATCACTCAATAAACCT 5485

Qy       1338 AAGAACCCTCATA 1350

Db       5484 AAGAACCCTCCTA 5472


RESULT 7
AP002956/c
LOCUS       AP002956               219574 bp    DNA     linear   PRI 15-AUG-2002
DEFINITION  Homo sapiens genomic DNA, chromosome 11q clone:R1105h09, complete
            sequence.
ACCESSION   AP002956
VERSION     AP002956.1  GI:22255355
KEYWORDS    HTG.
SOURCE      Homo sapiens DNA, clone:R1105h09.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Hattori,M., Toyoda,A., Taylor,T.D., Fujiyama,A., Yada,T.,
            Totoki,Y., Watanabe,H. and Sakaki,Y.
  TITLE     Homo sapiens 219,574 genomic DNA of 11q
  JOURNAL   Published Only in Database (2002)
REFERENCE   2  (bases 1 to 219574)
  AUTHORS   Hattori,M., Toyoda,A., Taylor,T.D., Fujiyama,A., Yada,T.,
            Totoki,Y., Watanabe,H. and Sakaki,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (24-NOV-2000) Masahira Hattori, The Institute of Physical
            and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
            1-7-22 Suehiro-chou,Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            (E-mail:hattori@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
            Tel:81-45-503-9111, Fax:81-45-503-9170)
COMMENT     Overlapping Clones: pD71A4, PQ38G8
            Cytogenetic Position: 21q22.1, region:D21S226-AML
            STS Markers (ePCR): stSG53747, SHGC-16045, D11S2450, SHGC-7143,
            WI-16056,
            stSG31054, sts-X57110, WI-8978, GDB:198117, stSG29284, SHGC-130657,
            D11S1941E, WI-14589, stSG60191, stSG50535, A002O48, stSG26946,
            stSG41092,
            TIGR-A005D38, Bdaa7g02, D11S869E, RH80030
            Contamination: none detected
            non-AGCT bases: none
            Additional author information
            Arai,Y., Kubo,T.,Ohki,M.
            National Cancer Center Research Institute
            Cancer Genomic Division
            5-1-1, Tsukiji, Chuo-ku, Tokyo, JAPAN
            zip: 104-0045
            phone: +81-3-3542-2511 ex4752, fax: +81-3-3542-0688 e-mail:
            yarai@ncc.go.jp.
FEATURES              Location/Qualifiers
     source           1. .219574
                      /organism="Homo sapiens"
                      /db_xref="taxon:9606"
                      /chromosome="11"
                      /map="11q"
                      /clone="R1105h09"
BASE COUNT    54883 a  50356 c  53844 g  60491 t
ORIGIN

  Query Match             66.1%;  Score 909.8;  DB 9;  Length 219574;
  Best Local Similarity   99.8%;  Pred. No. 2.6e-147;
  Matches  911;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy        438 CGGGACTGCCGGGACCTCGAGGGGACCCCGGGCCGCGAGGAGAGGCGGGACCCGCGGGGC 497

Db     209122 CAGGACTGCCGGGACCTCGAGGGGACCCCGGGCCGCGAGGAGAGGCGGGACCCGCGGGGC 209063

Qy        498 CCACCGGGCCTGCCGGGGAGTGCTCGGTGCCTCCGCGATCCGCCTTCAGCGCCAAGCGCT 557

Db     209062 CCACCGGGCCTGCCGGGGAGTGCTCGGTGCCTCCGCGATCCGCCTTCAGCGCCAAGCGCT 209003

Qy        558 CCGAGAGCCGGGTGCCTCCGCCGTCTGACGCACCCTTGCCCTTCGACCGCGTGCTGGTGA 617

Db     209002 CCGAGAGCCGGGTGCCTCCGCCGTCTGACGCACCCTTGCCCTTCGACCGCGTGCTGGTGA 208943

Qy        618 ACGAGCAGGGACATTACGACGCCGTCACCGGCAAGTTCACCTGCCAGGTGCCTGGGGTCT 677

Db     208942 ACGAGCAGGGACATTACGACGCCGTCACCGGCAAGTTCACCTGCCAGGTGCCTGGGGTCT 208883

Qy        678 ACTACTTCGCCGTCCATGCCACCGTCTACCGGGCCAGCCTGCAGTTTGATCTGGTGAAGA 737

Db     208882 ACTACTTCGCCGTCCATGCCACCGTCTACCGGGCCAGCCTGCAGTTTGATCTGGTGAAGA 208823

Qy        738 ATGGCGAATCCATTGCCTCTTTCTTCCAGTTTTTCGGGGGGTGGCCCAAGCCAGCCTCGC 797

Db     208822 ATGGCGAATCCATTGCCTCTTTCTTCCAGTTTTTCGGGGGGTGGCCCAAGCCAGCCTCGC 208763

Qy        798 TCTCGGGGGGGGCCATGGTGAGGCTGGAGCCTGAGGACCAAGTGTGGGTGCAGGTGGGTG 857

Db     208762 TCTCGGGGGGGGCCATGGTGAGGCTGGAGCCTGAGGACCAAGTGTGGGTGCAGGTGGGTG 208703

Qy        858 TGGGTGACTACATTGGCATCTATGCCAGCATCAAGACAGACAGCACCTTCTCCGGATTTC 917

Db     208702 TGGGTGACTACATTGGCATCTATGCCAGCATCAAGACAGACAGCACCTTCTCCGGATTTC 208643

Qy        918 TGGTGTACTCCGACTGGCACAGCTCCCCAGTCTTTGCTTAGTGCCCACTGCAAAGTGAGC 977

Db     208642 TGGTGTACTCCGACTGGCACAGCTCCCCAGTCTTTGCTTAGTGCCCACTGCAAAGTGAGC 208583

Qy        978 TCATGCTCTCACTCCTAGAAGGAGGGTGTGAGGCTGACAACCAGGTCATCCAGGAGGGCT 1037

Db     208582 TCATGCTCTCACTCCTAGAAGGAGGGTGTGAGGCTGACAACCAGGTCATCCAGGAGGGCT 208523

Qy       1038 GGCCCCCCTGGAATATTGTGAATGACTAGGGAGGTGGGGTAGAGCACTCTCCGTCCTGCT 1097

Db     208522 GGCCCCCCTGGAATATTGTGAATGACTAGGGAGGTGGGGTAGAGCACTCTCCGTCCTGCT 208463

Qy       1098 GCTGGCAAGGAATGGGAACAGTGGCTGTCTGCGATCAGGTCTGGCAGCATGGGGCAGTGG 1157

Db     208462 GCTGGCAAGGAATGGGAACAGTGGCTGTCTGCGATCAGGTCTGGCAGCATGGGGCAGTGG 208403
```

```
Qy       1158 CTGGATTTCTGCCCAAGACCAGAGGAGTGTGCTGTGCTGGCAAGTGTAAGTCCCCCAGTT 1217
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     208402 CTGGATTTCTGCCCAAGACCAGAGGAGTGTGCTGTGCTGGCAAGTGTAAGTCCCCCAGTT 208343

Qy       1218 GCTCTGGTCCAGGAGCCCACGGTGGGGTGCTCTCTTCCTGGTCCTCTGCTTCTCTGGATC 1277
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     208342 GCTCTGGTCCAGGAGCCCACGGTGGGGTGCTCTCTTCCTGGTCCTCTGCTTCTCTGGATC 208283

Qy       1278 CTCCCCACCCCCTCCTGCTCCTGGGGCCGGCCCTTTTCTCAGAGATCACTCAATAAACCT 1337
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     208282 CTCCCCACCCCCTCCTGCTCCTGGGGCCGGCCCTTTTCTCAGAGATCACTCAATAAACCT 208223

Qy       1338 AAGAACCCTCATA 1350
             ||||||||||| ||
Db     208222 AAGAACCCTCCTA 208210
```

RESULT 8
AP001557

```
LOCUS       AP001557               182429 bp    DNA     linear   HTG 03-NOV-2000
DEFINITION  Homo sapiens chromosome 11 clone RP11-680A7 map 11q23, WORKING
            DRAFT SEQUENCE, 21 unordered pieces.
ACCESSION   AP001557
VERSION     AP001557.3  GI:11094164
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      Homo sapiens DNA, clone:RP11-680A7.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 182429)
  AUTHORS   Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
            Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
  TITLE     Homo sapiens 182,429 genomic DNA of 11q23
  JOURNAL   Published Only in DataBase (2000)
REFERENCE   2  (bases 1 to 182429)
  AUTHORS   Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
            Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (29-MAR-2000) Masahira Hattori, The Institute of Physical
            and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
            1-7-22 Suehiro-chou, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            (E-mail:hattori@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
            Tel:81-45-503-9111, Fax:81-45-503-9170)
COMMENT     On Nov 3, 2000 this sequence version replaced gi:8117391.
            -------------- Genome Center
                Center: RIKEN Genomic Sciences Center(GSC)
                Center code: RIKEN
                Web site: http://hgp.gsc.riken.go.jp/
                Contact: hattori@gsc.riken.go.jp
            -------------- Project Information
                Center project name: HumDraft11
                Center clone name: RP11-680A7
            -------------- Summary Statistics
                Sequencing vector: PCR products; 100% of reads
                Chemistry: Dye-terminator ET-amersham; 100% of reads
                Assembly program: Phrap; version 0.990329
                Consensus quality: 175597 bases at least Q40
                Consensus quality: 178555 bases at least Q30
                Consensus quality: 179768 bases at least Q20
                Insert size: 180429; sum-of-contigs
                Quality coverage: 9.01x in Q20 bases; sum-of-contigs
            --------------
            NOTE: This is a 'working draft' sequence. It currently consists of
            21 contigs. The true order of the pieces is not known and their
            order in this sequence record is arbitrary. Gaps between the
            contigs are represented as runs N, but the exact sizes of the gaps
            are unknown. This record will be updated with the finished sequence
            as soon as it is available and the accession number will be
            preserved
                  1     30107 contig of  30107 bp in length
              30208     54357 contig of  24150 bp in length
              54458     69839 contig of  15382 bp in length
              69940     84048 contig of  14109 bp in length
              84149     95177 contig of  11029 bp in length
              95278    107763 contig of  12486 bp in length
             107864    119057 contig of  11194 bp in length
             119158    128134 contig of   8977 bp in length
             128235    136903 contig of   8669 bp in length
             137004    145887 contig of   8884 bp in length
             145988    153199 contig of   7212 bp in length
             153300    159565 contig of   6266 bp in length
             159666    163580 contig of   3915 bp in length
             163681    165152 contig of   1472 bp in length
             165253    166901 contig of   1649 bp in length
             167002    170460 contig of   3459 bp in length
             170561    172761 contig of   2201 bp in length
             172862    175605 contig of   2744 bp in length
             175706    178308 contig of   2603 bp in length
             178409    180918 contig of   2510 bp in length
             181019    182429 contig of   1411 bp in length.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 21 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *       1     30107: contig of 30107 bp in length
            *   30108 30207: gap of      100 bp
            *   30208    54357: contig of 24150 bp in length
            *   54358 54457: gap of      100 bp
            *   54458    69839: contig of 15382 bp in length
            *   69840 69939: gap of      100 bp
            *   69940    84048: contig of 14109 bp in length
            *   84049 84148: gap of      100 bp
            *   84149    95177: contig of 11029 bp in length
            *   95178 95277: gap of      100 bp
            *   95278   107763: contig of 12486 bp in length
            *  107764 107863: gap of      100 bp
            *  107864   119057: contig of 11194 bp in length
            *  119058 119157: gap of      100 bp
            *  119158   128134: contig of 8977 bp in length
            *  128135 128234: gap of      100 bp
            *  128235   136903: contig of 8669 bp in length
            *  136904 137003: gap of      100 bp
            *  137004   145887: contig of 8884 bp in length
            *  145888 145987: gap of      100 bp
            *  145988   153199: contig of 7212 bp in length
            *  153200 153299: gap of      100 bp
            *  153300   159565: contig of 6266 bp in length
            *  159566 159665: gap of      100 bp
            *  159666   163580: contig of 3915 bp in length
            *  163581 163680: gap of      100 bp
            *  163681   165152: contig of 1472 bp in length
            *  165153 165252: gap of      100 bp
            *  165253   166901: contig of 1649 bp in length
            *  166902 167001: gap of      100 bp
            *  167002   170460: contig of 3459 bp in length
            *  170461 170560: gap of      100 bp
            *  170561   172761: contig of 2201 bp in length
            *  172762 172861: gap of      100 bp
            *  172862   175605: contig of 2744 bp in length
            *  175606 175705: gap of      100 bp
            *  175706   178308: contig of 2603 bp in length
            *  178309 178408: gap of      100 bp
            *  178409   180918: contig of 2510 bp in length
            *  180919 181018: gap of      100 bp
            *  181019   182429: contig of 1411 bp in length.
FEATURES             Location/Qualifiers
     source          1. .182429
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="11"
                     /map="11q23"
                     /clone="RP11-680A7"
```

```
     misc_feature    1. .30107
                     /note="assembly_fragment"
     misc_feature    30208. .54357
                     /note="assembly_fragment"
     misc_feature    54458. .69839
                     /note="assembly_fragment"
     misc_feature    69940. .84048
                     /note="assembly_fragment"
     misc_feature    84149. .95177
                     /note="assembly_fragment"
     misc_feature    95278. .107763
                     /note="assembly_fragment"
     misc_feature    107864. .119057
                     /note="assembly_fragment"
     misc_feature    119158. .128134
                     /note="assembly_fragment"
     misc_feature    128235. .136903
                     /note="assembly_fragment"
     misc_feature    137004. .145887
                     /note="assembly_fragment"
     misc_feature    145988. .153199
                     /note="assembly_fragment"
     misc_feature    153300. .159565
                     /note="assembly_fragment"
     misc_feature    159666. .163580
                     /note="assembly_fragment"
     misc_feature    163681. .165152
                     /note="assembly_fragment clone_end:T7 vector_side:left"
     misc_feature    165253. .166901
                     /note="assembly_fragment clone_end:SP6 vector_side:left"
     misc_feature    167002. .170460
                     /note="assembly_fragment"
     misc_feature    170561. .172761
                     /note="assembly_fragment"
     misc_feature    172862. .175605
                     /note="assembly_fragment"
     misc_feature    175706. .178308
                     /note="assembly_fragment"
     misc_feature    178409. .180918
                     /note="assembly_fragment"
     misc_feature    181019. .182429
                     /note="assembly_fragment"
BASE COUNT    45724 a  42083 c  43875 g  48747 t   2000 others
ORIGIN

  Query Match             65.1%;  Score 896.2;  DB 2;  Length 182429;
  Best Local Similarity   99.6%;  Pred. No. 5.8e-145;
  Matches  909;  Conservative    0;  Mismatches    3;  Indels    1;  Gaps    1;

Qy       438 CGGGACTGCCGGGACCTCGAGGGGACCCCGGGCCGCGAGGAGAGGCGGGACCCGCGGGGC 497
            | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    108011 CAGGACTGCCGGGACCTCGAGGGGACCCCGGGCCGCGAGGAGAGGCGGGACCCGCGGGGC 108070

Qy       498 CCACCGGGCCTGCCGGGGAGTGCTCGGTGCCTCCGCGATCCGCCTTCAGCGCCAAGCGCT 557
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    108071 CCACCGGGCCTGCCGGGGAGTGCTCGGTGCCTCCGCGATCCGCCTTCAGCGCCAAGCGCT 108130

Qy       558 CCGAGAGCCGGGTGCCTCCGCCGTCTGACGCACCCTTGCCCTTCGACCGCGTGCTGGTGA 617
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    108131 CCGAGAGCCGGGTGCCTCCGCCGTCTGACGCACCCTTGCCCTTCGACCGCGTGCTGGTGA 108190

Qy       618 ACGAGCAGGGACATTACGACGCCGTCACCGGCAAGTTCACCTGCCAGGTGCCTGGGGTCT 677
            ||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    108191 ACGAGCA-GGACATTACGACGCCGTCACCGGCAAGTTCACCTGCCAGGTGCCTGGGGTCT 108249

Qy       678 ACTACTTCGCCGTCCATGCCACCGTCTACCGGGCCAGCCTGCAGTTTGATCTGGTGAAGA 737
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    108250 ACTACTTCGCCGTCCATGCCACCGTCTACCGGGCCAGCCTGCAGTTTGATCTGGTGAAGA 108309

Qy       738 ATGGCGAATCCATTGCCTCTTTCTTCCAGTTTTTCGGGGGGTGGCCCAAGCCAGCCTCGC 797
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    108310 ATGGCGAATCCATTGCCTCTTTCTTCCAGTTTTTCGGGGGGTGGCCCAAGCCAGCCTCGC 108369

Qy       798 TCTCGGGGGGGGCCATGGTGAGGCTGGAGCCTGAGGACCAAGTGTGGGTGCAGGTGGGTG 857
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    108370 TCTCGGGGGGGGCCATGGTGAGGCTGGAGCCTGAGGACCAAGTGTGGGTGCAGGTGGGTG 108429

Qy       858 TGGGTGACTACATTGGCATCTATGCCAGCATCAAGACAGACAGCACCTTCTCCGGATTTC 917
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    108430 TGGGTGACTACATTGGCATCTATGCCAGCATCAAGACAGACAGCACCTTCTCCGGATTTC 108489

Qy       918 TGGTGTACTCCGACTGGCACAGCTCCCCAGTCTTTGCTTAGTGCCCACTGCAAAGTGAGC 977
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    108490 TGGTGTACTCCGACTGGCACAGCTCCCCAGTCTTTGCTTAGTGCCCACTGCAAAGTGAGC 108549

Qy       978 TCATGCTCTCACTCCTAGAAGGAGGGTGTGAGGCTGACAACCAGGTCATCCAGGAGGGCT 1037
            |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db    108550 TCATGCTCTCACTCCTAGAAGGAGGGTGTGAGGCTGACAACCTGGTCATCCAGGAGGGCT 108609

Qy      1038 GGCCCCCCTGGAATATTGTGAATGACTAGGGAGGTGGGGTAGAGCACTCTCCGTCCTGCT 1097
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    108610 GGCCCCCCTGGAATATTGTGAATGACTAGGGAGGTGGGGTAGAGCACTCTCCGTCCTGCT 108669

Qy      1098 GCTGGCAAGGAATGGGAACAGTGGCTGTCTGCGATCAGGTCTGGCAGCATGGGGCAGTGG 1157
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    108670 GCTGGCAAGGAATGGGAACAGTGGCTGTCTGCGATCAGGTCTGGCAGCATGGGGCAGTGG 108729

Qy      1158 CTGGATTTCTGCCCAAGACCAGAGGAGTGTGCTGTGCTGGCAAGTGTAAGTCCCCCAGTT 1217
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    108730 CTGGATTTCTGCCCAAGACCAGAGGAGTGTGCTGTGCTGGCAAGTGTAAGTCCCCCAGTT 108789

Qy      1218 GCTCTGGTCCAGGAGCCCACGGTGGGGTGCTCTCTTCCTGGTCCTCTGCTTCTCTGGATC 1277
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    108790 GCTCTGGTCCAGGAGCCCACGGTGGGGTGCTCTCTTCCTGGTCCTCTGCTTCTCTGGATC 108849

Qy      1278 CTCCCCACCCCCTCCTGCTCCTGGGGCCGGCCCTTTTCTCAGAGATCACTCAATAAACCT 1337
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    108850 CTCCCCACCCCCTCCTGCTCCTGGGGCCGGCCCTTTTCTCAGAGATCACTCAATAAACCT 108909

Qy      1338 AAGAACCCTCATA 1350
            |||||||||| ||
Db    108910 AAGAACCCTCCTA 108922


RESULT 9
AF469650
LOCUS       AF469650                4220 bp    mRNA    linear   ROD 31-JUL-2002
DEFINITION  Mus musculus membrane-type frizzled-related protein (Mfrp) and
            complement-c1q tumor necrosis factor-related protein (C1qtnf5)
            mRNAs, complete cds.
ACCESSION   AF469650
VERSION     AF469650.1  GI:22023766
KEYWORDS    .
SOURCE      Mus musculus.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 4220)
  AUTHORS   Kameya,S., Hawes,N.L., Chang,B., Heckenlively,J.R., Naggert,J.K.
            and Nishina,P.M.
  TITLE     Mfrp, a gene encoding a frizzled related protein, is mutated in the
            mouse retinal degeneration 6
  JOURNAL   Hum. Mol. Genet. 11 (16), 1879-1886 (2002)
   PUBMED   12140190
REFERENCE   2  (bases 1 to 4220)
  AUTHORS   Kameya,S., Naggert,J.K. and Nishina,P.M.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-JAN-2002) The Jackson Laboratory, 600 Main Street,
            Bar Harbor, ME 04609, USA
FEATURES             Location/Qualifiers
     source          1. .4220
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /chromosome="9"
                     /map="26 cM"
```

```
     gene            52. .1806
                     /gene="Mfrp"
     CDS             52. .1806
                     /gene="Mfrp"
                     /note="MFRP"
                     /codon_start=1
                     /product="membrane-type frizzled-related protein"
                     /protein_id="AAM89216.1"
                     /db_xref="GI:22023767"
                     /translation="MKDYDDVILRPEASELSKTEFCNPAFDPEAGPSCPPPALQRDVG
                     SRLQAPWHAQRLRGLQPDCHFSWFCILLLSGLLLLLLGLLVAVILAQLQATSLPRTTK
                     NPLLTRGLTPMGVIPSTTPNTTTTTTTTTPARTGQQEAAMSPTHQTTCGGLLPGPSGF
                     FSSPNYPDLYPPLSHCVWHIQVAAGQTIQLKIQALSIESMLTCLFDRLEIISEPTGPL
                     LRVCGKTPPATLNTNTSHLRVSFVSDNDVEGSGFQAWYQAVAPGHWSCAHNEFHCDLL
                     LCLKRDSVCDGITECADGSDEANCSAKTLGCGGNLTGLYGVFSTPNYPQHYPHQQLCT
                     WYIEVPVGYGIRLEFHNFSLEAQAECKFDYVEVYEASNLGTFSFLGRFCGAEPPLNVV
                     SSMHQLAVIFKTDLGISSGGFLATYQAINTTESGCPWAEFCQSGGYRDLQWMCDLWKD
                     CANDSNDNCSSHLSPQPDLTCEPVQVEMCLGLSYNTTAFPNIWVGLATQTEVTDILRG
                     YKSLTSLPCYQTFQRFLCGLLVPRCTSLGTILPPCRSVCQAAEQQCQSSLALLGTPWP
                     FNCNRLPVAASLEACSQP"
     gene            3097. .3828
                     /gene="C1qtnf5"
     CDS             3097. .3828
                     /gene="C1qtnf5"
                     /note="C1QTNF5"
                     /codon_start=1
                     /product="complement-c1q tumor necrosis factor-related
                     protein"
                     /protein_id="AAM89217.1"
                     /db_xref="GI:22023768"
                     /translation="MRPLLALLLLGLVSGSPPLDDNKIPSLCPGQPGLPGTPGHHGSQ
                     GLPGRDGRDGRDGAPGAPGEKGEGGRPGLPGPRGEPGPRGEAGPMGAIGPAGECSVPP
                     RSAFSAKRSESRVPPPADTPLPFDRVLLNEQGHYDPTTGKFTCQVPGVYYFAVHATVY
                     RASLQFDLVKNGQSIASFFQYFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIY
                     ASIKTDSTFSGFLVYSDWHSSPVFA"
BASE COUNT      873 a    1282 c    1147 g     918 t
ORIGIN
```

```
  Query Match           50.6%;  Score 696.2;  DB 10;  Length 4220;
  Best Local Similarity  78.0%;  Pred. No. 2.3e-110;
  Matches  921;  Conservative    0;  Mismatches  238;  Indels   22;  Gaps    6;

Qy       189 GAGAGGACCCCCGGCGTCCGGGCTCCCGGTGCCAGCGCTATGAGGCCACTCCTCGTCCTGC 248
            ||||||  || || ||| | ||  |||| | || | | |||||||||| || | ||| |
Db      3059 GAGAGGTTCCTGGAGTCTGAGCCTCCGGGGTCACCACCATGAGGCCACTTCTTGCCCTTC 3118

Qy       249 TGCTCCTGGGCCTGGCGGCCGGCTCGCCCCCACTGGACGACAACAAGATCCCCAGCCTCT 308
            |||| ||||| |||| | | ||||| || || ||||||||||||||||||||||||||| |
Db      3119 TGCTTCTGGGTCTGGTGTCAGGCTCTCCTCCTCTGGACGACAACAAGATCCCCAGCCTGT 3178

Qy       309 GCCCGGGGCACCCCGGCCTTCCAGGCACGCCGGGCCACCATGGCAGCCAGGGCTTGCCGG 368
            | || ||||| |||||||||||||||| || || |||||||||||||| ||| |||| |
Db      3179 GTCCCGGGCAGCCCGGCCTTCCAGGCACACCAGGTCACCATGGCAGCCAAGGCCTGCCTG 3238

Qy       369 GCCGCGATGGCCGCGACGGCCGCGACGGCGCGCCCGGGGCTCCGGGAGAGAAAGGCGAGG 428
            |||| || ||||| || |||||||||| || ||||| ||||||||||||||||||||||
Db      3239 GCCGTGACGGCCGTGATGGCCGCGACGGTGCACCCGGAGCTCCGGGAGAGAAAGGCGAGG 3298

Qy       429 GCGGGAGGCCGGGACTGCCGGGACCTCGAGGGGACCCCGGGCCGCGAGGAGAGGCGGGAC 488
            ||||||| |||||||| || || || || ||||| |||||||||| |||||||| || |
Db      3299 GCGGGAGACCGGGACTACCTGGCCCACGTGGGGAGCCCGGGCCGCGTGGAGAGGCAGGGC 3358

Qy       489 CCGCGGGGCCCACCGGGCCTGCCGGGGAGTGCTCGGTGCCTCCGCGATCCGCCTTCAGCG 548
            ||  |||| | | ||||||||| |||||||||||||| || || ||||| |||||||| |
Db      3359 CCATGGGGGCTATCGGGCCTGCGGGGGAGTGCTCGGTACCCCCACGATCAGCCTTCAGTG 3418

Qy       549 CCAAGCGCTCCGAGAGCCGGGTGCCTCCGCCGTCTGACGCACCCTTGCCCTTCGACCGCG 608
            ||||||| |||||||||||||| |||||||| | ||| ||||| | || |||||||| |
Db      3419 CCAAGCGATCCGAGAGCCGGGTACCTCCGCCAGCCGACACACCCCTACCTTTCGACCGTG 3478

Qy       609 TGCTGGTGAACGAGCAGGGACATTACGACGCCGTCACCGGCAAGTTCACCTGCCAGGTGC 668
            ||||| | || |||||||| ||||||||| ||    || |||||||||||||||| ||||
Db      3479 TGCTGCTAAATGAGCAGGGCCATTACGACCCCACTACTGGCAAGTTCACCTGCCAAGTGC 3538

Qy       669 CTGGGGTCTACTACTTCGCCGTCCATGCCACCGTCTACCGGGCCAGCCTGCAGTTTGATC 728
            |||| |||||||||||| || || |||| |||||||||||||||| ||||||||||||
Db      3539 CTGGCGTCTACTACTTTGCTGTGCACGCCACTGTCTACCGGGCCAGCTTGCAGTTTGATC 3598

Qy       729 TGGTGAAGAATGGCGAATCCATTGCCTCTTTCTTCCAGTTTTTCGGGGGGTGGCCCAAGC 788
            | || || || ||  | ||||| |||||||||||||||| ||| ||||||||||||||||
Db      3599 TTGTCAAAAACGGGCAGTCCATCGCCTCTTTCTTCCAGTATTTTGGGGGGTGGCCCAAGC 3658

Qy       789 CAGCCTCGCTCTCGGGGGGGGCCATGGTGAGGCTGGAGCCTGAGGACCAAGTGTGGGTGC 848
            |||||||||||| ||||| || ||||| ||||| || |||||||||| ||||||||||
Db      3659 CAGCCTCGCTCTCAGGGGGTGCGATGGTAAGGCTAGAACCTGAGGACCAGGTGTGGGTGC 3718

Qy       849 AGGTGGGTGTGGGTGACTACATTGGCATCTATGCCAGCATCAAGACAGACAGCACCTTCT 908
            ||||||| |||||||| ||||||||||||||||||||||||||||||||||||| |||||||
Db      3719 AGGTGGGCGTGGGTGATTACATTGGCATCTATGCCAGCATCAAGACAGACAGTACCTTCT 3778

Qy       909 CCGGATTTCTGGTGTACTCCGACTGGCACAGCTCCCCAGTCTTTGCTTAGTGCCCACTGC 968
            | |||||||| || || || |||||||||||||||||||||||| |||||   | || ||
Db      3779 CTGGATTTCTCGTCTATTCTGACTGGCACAGCTCCCCAGTCTTCGCTTAAAACACAGTGA 3838

Qy       969 AAAGTGAGCTCATGCTCTCACTCCTAGAAGGAGGGTGTGAGGCTGACAACCAGGTCATCC 1028
            |    |||||    ||  | |  |       | | |||  |     |   ||   ||| |
Db      3839 ACCCGGAGCTGGCACTTGCTCCTCAGTGGAGGGTGTGACACTAACCCGCGCAACGCATAC 3898

Qy      1029 AGGAGGGCTGGCCCCCCTGGAATATTGTGAATGACTAGGGAGGTGGGGTAGAGCACTCTC 1088
               |  ||   | |||||||||||||||||||||||||  ||| | | || ||  ||||  |
Db      3899 CAGGAGGGCTGGCCCCCTGGAATATTGTGAATGACTTAGGAAGAGAGGGAG-CCACTTCC 3957

Qy      1089 CGTCCTGCTGCTGGCAAGGAATGGGAACAGTGGCTGTCTGCGATCAGGTCTGGCAGCATG 1148
              ||||  |||||||||| |||||  |||  |||||||| | |||     | |||| ||
Db      3958 AGTCCCACTGCTGGCAATGAATGGAGACA--GGCTGTCTGAGGTCA----AGACAGCGTG 4011

Qy      1149 GGGCAGTGGCTGGATTTCTGCCCAAGACCAGAGGAGTGTGCTGTGCTGGCAAGTGTAAGT 1208
            | |||||||||| |||||||||| |||   | || ||   |  |||||||  |||  ||
Db      4012 GAGCAGTGGCTGGGTTTCTGCCCAGGACTTTA-GAATGCAGTAGGCTGGCAGCTGTGGGT 4070

Qy      1209 CCCCCAGTTGCTCTGGTCCAGGAGCCCACGGTGGGGTGCTCTCTTCCTGGTCCTCTG--- 1265
            |           |||| ||||||  ||| ||||||  |||||  ||||| ||||| ||
Db      4071 C----------CTGGCCCAGGACTCCAAGGTGGGATGCTCCATTCCTAGTCCTGTGTCC 4119

Qy      1266 CTTCTCTGGATCCTCCCCACCCCCTCCTGCTCCTGGGGCCGGCCCTTTTCTCAGAGATCA 1325
            | |||  |   |   | |  | || |||||||  |||| |||| ||||||||||| |||
Db      4120 CCTCTAGGTCCCTGACTCCATCTCTGCTGCTCCCAGGGCAGGCCTTTTTCTCAGAGGTCA 4179

Qy      1326 CTCAATAAACCTAAGAACCCTCATAAAAAAAAAAAAAAAAA 1366
            || |||||||||| | ||  |  ||||||||||||||||
Db      4180 CTTAATAAACCTAAAATCCTCCCAAAAAAAAAAAAAAAAA 4220
```

```
RESULT 10
AP001003
LOCUS       AP001003               173038 bp    DNA     linear   HTG 30-MAY-2000
DEFINITION  Homo sapiens chromosome 11 clone RP11-680C5 map 11q23, WORKING
            DRAFT SEQUENCE, 28 unordered pieces.
ACCESSION   AP001003
VERSION     AP001003.2  GI:8117674
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      Homo sapiens DNA, clone:RP11-680C5.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 173038)
  AUTHORS   Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
            Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
  TITLE     Homo sapiens 173,038 genomic DNA of 11q23
  JOURNAL   Published Only in DataBase (2000)
REFERENCE   2  (bases 1 to 173038)
  AUTHORS   Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
            Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (05-JAN-2000) Masahira Hattori, The Institute of Physical
            and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
```

```
            Kitasato Univ., 1-15-1 Kitasato, Sagamihara, Kanagawa 228-8555,
            Japan (E-mail:hattori@gsc.riken.go.jp,
            URL:http://hgp.gsc.riken.go.jp/, Tel:81-42-778-9923,
            Fax:81-42-778-9924)
COMMENT     On May 31, 2000 this sequence version replaced gi:6997762.
            -------------- Genome Center
                Center: RIKEN Genomic Sciences Center(GSC)
                Center code: RIKEN
                Web site: http://hgp.gsc.riken.go.jp/
                Contact: hattori@gsc.riken.go.jp
            -------------- Project Information
                Center project name: HumDraft11
                Center clone name: RP11-680C5
            -------------- Summary Statistics
                Sequencing vector: PCR products; 100% of reads
                Chemistry: Dye-terminator ET-amersham; 100% of reads
                Assembly program: Phrap; version 0.990329
                Consensus quality: 149676 bases at least Q40
                Consensus quality: 160868 bases at least Q30
                Consensus quality: 167161 bases at least Q20
                Insert size: 170338; sum-of-contigs
                Quality coverage: 4.33x in Q20 bases; sum-of-contigs
            --------------.
            NOTE: This is a 'working draft' sequence. It currently consists of
            28 contigs. The true order of the pieces is not known and their
            order in this sequence record is arbitrary. Gaps between the
            contigs are represented as runs N, but the exact sizes of the gaps
            are unknown. This record will be updated with the finished sequence
            as soon as it is available and the accession number will be
            preserved
                  1     15558 contig of  15558 bp in length
              15659     33689 contig of  18031 bp in length
              33790     47432 contig of  13643 bp in length
              47533     63392 contig of  15860 bp in length
              63493     74234 contig of  10742 bp in length
              74335     85357 contig of  11023 bp in length
              85458     93597 contig of   8140 bp in length
              93698    101983 contig of   8286 bp in length
             102084    110022 contig of   7939 bp in length
             110123    118116 contig of   7994 bp in length
             118217    126464 contig of   8248 bp in length
             126565    132136 contig of   5572 bp in length
             132237    135290 contig of   3054 bp in length
             135391    140229 contig of   4839 bp in length
             140330    143896 contig of   3567 bp in length
             143997    147572 contig of   3576 bp in length
             147673    150446 contig of   2774 bp in length
             150547    153715 contig of   3169 bp in length
             153816    157273 contig of   3458 bp in length
             157374    159817 contig of   2444 bp in length
             159918    161743 contig of   1826 bp in length
             161844    163684 contig of   1841 bp in length
             163785    165216 contig of   1432 bp in length
             165317    166657 contig of   1341 bp in length
             166758    168307 contig of   1550 bp in length
             168408    170030 contig of   1623 bp in length
             170131    171725 contig of   1595 bp in length
             171826    173038 contig of   1213 bp in length
            Sequence updated (26-May-2000).
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 28 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *       1    15558: contig of 15558 bp in length
            *   15559 15658: gap of      100 bp
            *   15659    33689: contig of 18031 bp in length
            *   33690 33789: gap of      100 bp
            *   33790    47432: contig of 13643 bp in length
            *   47433 47532: gap of      100 bp
            *   47533    63392: contig of 15860 bp in length
            *   63393 63492: gap of      100 bp
            *   63493    74234: contig of 10742 bp in length
            *   74235 74334: gap of      100 bp
            *   74335    85357: contig of 11023 bp in length
            *   85358 85457: gap of      100 bp
            *   85458    93597: contig of 8140 bp in length
            *   93598 93697: gap of      100 bp
            *   93698   101983: contig of 8286 bp in length
            *  101984 102083: gap of      100 bp
            *  102084   110022: contig of 7939 bp in length
            *  110023 110122: gap of      100 bp
            *  110123   118116: contig of 7994 bp in length
            *  118117 118216: gap of      100 bp
            *  118217   126464: contig of 8248 bp in length
            *  126465 126564: gap of      100 bp
            *  126565   132136: contig of 5572 bp in length
            *  132137 132236: gap of      100 bp
            *  132237   135290: contig of 3054 bp in length
            *  135291 135390: gap of      100 bp
            *  135391   140229: contig of 4839 bp in length
            *  140230 140329: gap of      100 bp
            *  140330   143896: contig of 3567 bp in length
            *  143897 143996: gap of      100 bp
            *  143997   147572: contig of 3576 bp in length
            *  147573 147672: gap of      100 bp
            *  147673   150446: contig of 2774 bp in length
            *  150447 150546: gap of      100 bp
            *  150547   153715: contig of 3169 bp in length
            *  153716 153815: gap of      100 bp
            *  153816   157273: contig of 3458 bp in length
            *  157274 157373: gap of      100 bp
            *  157374   159817: contig of 2444 bp in length
            *  159818 159917: gap of      100 bp
            *  159918   161743: contig of 1826 bp in length
            *  161744 161843: gap of      100 bp
            *  161844   163684: contig of 1841 bp in length
            *  163685 163784: gap of      100 bp
            *  163785   165216: contig of 1432 bp in length
            *  165217 165316: gap of      100 bp
            *  165317   166657: contig of 1341 bp in length
            *  166658 166757: gap of      100 bp
            *  166758   168307: contig of 1550 bp in length
            *  168308 168407: gap of      100 bp
            *  168408   170030: contig of 1623 bp in length
            *  170031 170130: gap of      100 bp
            *  170131   171725: contig of 1595 bp in length
            *  171726 171825: gap of      100 bp
            *  171826   173038: contig of 1213 bp in length.
FEATURES                Location/Qualifiers
     source             1. .173038
                        /organism="Homo sapiens"
                        /db_xref="taxon:9606"
                        /chromosome="11"
                        /map="11q23"
                        /clone="RP11-680C5"
     misc_feature       1. .15558
                        /note="assembly_fragment"
     misc_feature       15659. .33689
                        /note="assembly_fragment"
     misc_feature       33790. .47432
                        /note="assembly_fragment"
     misc_feature       47533. .63392
                        /note="assembly_fragment"
     misc_feature       63493. .74234
                        /note="assembly_fragment"
     misc_feature       74335. .85357
                        /note="assembly_fragment"
     misc_feature       85458. .93597
                        /note="assembly_fragment"
     misc_feature       93698. .101983
                        /note="assembly_fragment"
     misc_feature       102084. .110022
```

```
                     /note="assembly_fragment"
     misc_feature    110123. .118116
                     /note="assembly_fragment"
     misc_feature    118217. .126464
                     /note="assembly_fragment"
     misc_feature    126565. .132136
                     /note="assembly_fragment"
     misc_feature    132237. .135290
                     /note="assembly_fragment"
     misc_feature    135391. .140229
                     /note="assembly_fragment"
     misc_feature    140330. .143896
                     /note="assembly_fragment"
     misc_feature    143997. .147572
                     /note="assembly_fragment"
     misc_feature    147673. .150446
                     /note="assembly_fragment clone_end:T7 vector_side:right"
     misc_feature    150547. .153715
                     /note="assembly_fragment"
     misc_feature    153816. .157273
                     /note="assembly_fragment"
     misc_feature    157374. .159817
                     /note="assembly_fragment"
     misc_feature    159918. .161743
                     /note="assembly_fragment"
     misc_feature    161844. .163684
                     /note="assembly_fragment"
     misc_feature    163785. .165216
                     /note="assembly_fragment clone_end:SP6 vector_side:left"
     misc_feature    165317. .166657
                     /note="assembly_fragment"
     misc_feature    166758. .168307
                     /note="assembly_fragment"

  Query Match             50.4%;  Score 694;  DB 2;  Length 173038;
  Best Local Similarity   98.8%;  Pred. No. 4e-110;
  Matches  741;  Conservative    0;  Mismatches    5;  Indels    4;  Gaps    4;

Qy        604 CCGCGTGCTGGTGAACGAGCAGGGACATTACGACGCCGTCACCGGCAAGTTCACCTGCCA 663
              |||||||||||||||||||| |||||||||||||||||||||||||||||||||||||||
Db     166758 CCGCGTGCTGGTGAACGAGCA-GGACATTACGACGCCGTCACCGGCAAGTTCACCTGCCA 166816

Qy        664 GGTGCCTGGGGTCTACTACTTCGCCGTCCATGCCACCGTCTACCGGGCCAGCCTGCAGTT 723
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     166817 GGTGCCTGGGGTCTACTACTTCGCCGTCCATGCCACCGTCTACCGGGCCAGCCTGCAGTT 166876

Qy        724 TGATCTGGTGAAGAATGGCGAATCCATTGCCTCTTTCTTCCAGTTTTTC-GGGGGGTGGC 782
              |||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db     166877 TGATCTGGTGAAGAATGGCGAATCCATTGCCTCTTTCTTCCAGTTTTTCGGGGGGGTGGC 166936

Qy        783 CCAAGCCAGCCTCGCTCTCGGGGGGGGCCATGGTGAGGCTGGAGCCTGAGGACCAAGTGT 842
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     166937 CCAAGCCAGCCTCGCTCTCGGGGGGGGCCATGGTGAGGCTGGAGCCTGAGGACCAAGTGT 166996

Qy        843 GGGTGCAGGTGGGTGTGGGTGACTACATTGGCATCTATGCCAGCATCAAGACAGACAGCA 902
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     166997 GGGTGCAGGTGGGTGTGGGTGACTACATTGGCATCTATGCCAGCATCAAGACAGACAGCA 167056

Qy        903 CCTTCTCCGGATTTCTGGTGTACTCCGACTGGCACAGCTCCCCAGTCTTTGCTTAGTGCC 962
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     167057 CCTTCTCCGGATTTCTGGTGTACTCCGACTGGCACAGCTCCCCAGTCTTTGCTTAGTGCC 167116

Qy        963 CACTGCAAAGTGAGCTCATGCTCTCACTCCTAGAAGGAGGGTGTGAGGCTGACAACCAGG 1022
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db     167117 CACTGCAAAGTGAGCTCATGCTCTCACTCCTAGAAGGAGGGTGTGAGGCTGACAACCTGG 167176

Qy       1023 TCAT-CCAGGAGGGCTGGCCCCCC-TGGAATATTGTGAATGACTAGGGAGGTGGGGTAGA 1080
              |||| |||||||||||||||||| |||||||||||| |||||||||||  |||||||
Db     167177 TCATCCCAGGAGGGCTGGCCCCCCTTGGAATATTGTGACTGACTAGGGAGGGTGGGTAGA 167236

Qy       1081 GCACTCTCCGTCCTGCTGCTGGCAAGGAATGGGAACAGTGGCTGTCTGCGATCAGGTCTG 1140
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     167237 GCACTCTCCGTCCTGCTGCTGGCAAGGAATGGGAACAGTGGCTGTCTGCGATCAGGTCTG 167296

Qy       1141 GCAGCATGGGGCAGTGGCTGGATTTCTGCCCAAGACCAGAGGAGTGTGCTGTGCTGGCAA 1200
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     167297 GCAGCATGGGGCAGTGGCTGGATTTCTGCCCAAGACCAGAGGAGTGTGCTGTGCTGGCAA 167356

Qy       1201 GTGTAAGTCCCCCAGTTGCTCTGGTCCAGGAGCCCACGGTGGGGTGCTCTCTTCCTGGTC 1260
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     167357 GTGTAAGTCCCCCAGTTGCTCTGGTCCAGGAGCCCACGGTGGGGTGCTCTCTTCCTGGTC 167416

Qy       1261 CTCTGCTTCTCTGGATCCTCCCCACCCCCTCCTGCTCCTGGGGCCGGCCCTTTTCTCAGA 1320
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     167417 CTCTGCTTCTCTGGATCCTCCCCACCCCCTCCTGCTCCTGGGGCCGGCCCTTTTCTCAGA 167476

Qy       1321 GATCACTCAATAAACCTAAGAACCCTCATA 1350
              |||||||||||||||||||||||||||| ||
Db     167477 GATCACTCAATAAACCTAAGAACCCTCCTA 167506


RESULT 11
BC025174
LOCUS       BC025174                1234 bp    mRNA    linear   ROD 07-AUG-2002
DEFINITION  Mus musculus, clone MGC:36714 IMAGE:3978387, mRNA, complete cds.
ACCESSION   BC025174
VERSION     BC025174.1  GI:19263583
KEYWORDS    MGC.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 1234)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (05-MAR-2002) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.ncl.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Gilbert Smith, Ph.D.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Baylor College of Medicine Human Genome
            Sequencing Center
            Center code: BCM-HGSC
            Web site: http://www.hgsc.bcm.tmc.edu/cdna/
            Contact: amg@bcm.tmc.edu
            Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Hale, S.M.,
            Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M.,
            Richards, S., Gibbs, R.A.

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAK Plate: 61 Row: f Column: 24
            This clone was selected for full length sequencing because it
            passed the following selection criteria: Hexamer frequency ORF
            analysis.
FEATURES             Location/Qualifiers
     source          1. .1234
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /map="FVB/N"
                     /clone="MGC:36714 IMAGE:3978387"
                     /tissue_type="Mammary tumor. Metallothionien-TGF alpha
                     model. 10 month old virgin mouse. Taken by biopsy."
                     /clone_lib="NCI_CGAP_Mam1"
                     /lab_host="DH10B"
                     /note="Vector: pCMV-SPORT6"
     CDS             112. .843
                     /codon_start=1
                     /product="Unknown (protein for MGC:36714)"
                     /protein_id="AAH25174.1"
                     /db_xref="GI:19263584"
```

```
                     /translation="MRPLLALLLLGLVSGSPPLDDNKIPSLCPGQPGLPGTPGHHGSQ
                     GLPGRDGRDGRDGAPGAPGEKGEGGRPGLPGPRGEPGPRGEAGPMGAIGPAGECSVPP
                     RSAFSAKRSESRVPPPADTPLPFDRVLLNEQGHFDPTTGKFTCQVPGVYYFAVHATVY
                     RASLQFDLVKNGQSIASFFQYFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIY
                     ASIKTDSTFSGFLVYSDWHSSPVFA"
BASE COUNT       257 a    365 c    369 g    243 t
ORIGIN

  Query Match           50.4%;  Score 693.6;  DB 10;  Length 1234;
  Best Local Similarity  77.9%;  Pred. No. 7e-110;
  Matches  919;  Conservative    0;  Mismatches  239;  Indels   22;  Gaps    6;

Qy       189 GAGAGGACCCCGGCGTCCGGGCTCCCGGTGCCAGCGCTATGAGGCCACTCCTCGTCCTGC 248
Db        74 GAGAGGTTCCTGGAGTCTGAGCCTCCGGGGTCACCACCATGAGGCCACTTCTTGCCCTTC 133

Qy       249 TGCTCCTGGGCCTGGCGGCCGGCTCGCCCCCACTGGACGACAACAAGATCCCCAGCCTCT 308
Db       134 TGCTTCTGGGTCTGGTGTCAGGCTCTCCTCCTCTGGACGACAACAAGATCCCCAGCCTGT 193

Qy       309 GCCCGGGGCACCCCGGCCTTCCAGGCACGCCGGGCCACCATGGCAGCCAGGGCTTGCCGG 368
Db       194 GTCCCGGGCAGCCCGGCCTTCCAGGCACACCAGGTCACCATGGCAGCCAAGGCCTGCCTG 253

Qy       369 GCCGCGATGGCCGCGACGGCCGCGACGGCGCGCCCGGGGCTCCGGGAGAGAAAGGCGAGG 428
Db       254 GCCGTGACGGCCGTGATGGCCGCGACGGTGCACCCGGAGCCCCGGGAGAGAAAGGCGAGG 313

Qy       429 GCGGGAGGCCGGGACTGCCGGGACCTCGAGGGGACCCCGGGCCGCGAGGAGAGGCGGGAC 488
Db       314 GCGGGAGACCGGGACTACCTGGGCCACGTGGGGAGCCCGGGCCGCGTGGAGAGGCAGGGC 373

Qy       489 CCGCGGGGCCCACCGGGCCTGCCGGGGAGTGCTCGGTGCCTCCGCGATCCGCCTTCAGCG 548
Db       374 CCATGGGGGCTATCGGGCCTGCGGGGGAGTGCTCGGTACCCCCACGATCAGCCTTCAGTG 433

Qy       549 CCAAGCGCTCCGAGAGCCGGGTGCCTCCGCCGTCTGACGCACCCTTGCCCTTCGACCGCG 608
Db       434 CCAAGCGATCCGAGAGCCGGGTACCTCCGCCAGCCGACACACCCCTACCTTTCGACCGTG 493

Qy       609 TGCTGGTGAACGAGCAGGGACATTACGACGCCGTCACCGGCAAGTTCACCTGCCAGGTGC 668
Db       494 TGCTGCTAAATGAGCAGGGCCATTTCGACCCCACTACTGGCAAGTTCACCTGCCAAGTGC 553

Qy       669 CTGGGGTCTACTACTTCGCCGTCCATGCCACCGTCTACCGGGCCAGCCTGCAGTTTGATC 728
Db       554 CTGGCGTCTACTACTTTGCTGTGCACGCCACTGTCTACCGGGCCAGCTTGCAGTTTGATC 613

Qy       729 TGGTGAAGAATGGCGAATCCATTGCCTCTTTCTTCCAGTTTTTCGGGGGGTGGCCCAAGC 788
Db       614 TTGTCAAAAACGGGCAGTCCATCGCCTCTTTCTTCCAGTATTTTGGGGGGTGGCCCAAGC 673

Qy       789 CAGCCTCGCTCTCGGGGGGGGCCATGGTGAGGCTGGAGCCTGAGGACCAAGTGTGGGTGC 848
Db       674 CAGCCTCGCTCTCAGGGGGTGCGATGGTAAGGCTAGAACCTGAGGACCAGGTGTGGGTGC 733

Qy       849 AGGTGGGTGTGGGTGACTACATTGGCATCTATGCCAGCATCAAGACAGACAGCACCTTCT 908
Db       734 AGGTGGGCGTGGGTGATTACATTGGCATCTATGCCAGCATCAAGACAGACAGTACCTTCT 793

Qy       909 CCGGATTTCTGGTGTACTCCGACTGGCACAGCTCCCCAGTCTTTGCTTAGTGCCCACTGC 968
Db       794 CTGGATTTCTCGTCTATTCTGACTGGCACAGCTCCCCAGTCTTCGCTTAAAACACAGTGA 853

Qy       969 AAAGTGAGCTCATGCTCTCACTCCTAGAAGGAGGGTGTGAGGCTGACAACCAGGTCATCC 1028
Db       854 ACCCGGAGCTGGCACTTGCTCCTCAGTGGAGGGTGTGACACTAACCCGCGCAGCGCATAC 913

Qy      1029 AGGAGGGCTGGCCCCCCTGGAATATTGTGAATGACTAGGGAGGTGGGGTAGAGCACTCTC 1088
Db       914 CAGGAGGGCTGGCCCCCTGGAATATTGTGAATGACTTAGGAAGAGAGGGAG-CCACTTCC 972

Qy      1089 CGTCCTGCTGCTGGCAAGGAATGGGAACAGTGGCTGTCTGCGATCAGGTCTGGCAGCATG 1148
Db       973 GGTCCCACTGCTGGCAATGAATGGAGACA--GGCTGTCTGAGGTCA----AGACAGCGTG 1026

Qy      1149 GGGCAGTGGCTGGATTTCTGCCCAAGACCAGAGGAGTGTGCTGTGCTGGCAAGTGTAAGT 1208
Db      1027 GAGCAGTGGCTGGGTTTCTGCCCAGGACTTTA-GAATGCAGTAGGCTGGCAGCTGTGGGT 1085

Qy      1209 CCCCCAGTTGCTCTGGTCCAGGAGCCCACGGTGGGGTGCTCTCTTCCTGGTCCTCTG--- 1265
Db      1086 C----------CTGGCCCAGGACTCCAAGGTGGGATGCTCCATTCCTAGTCCTGTGCCC 1134

Qy      1266 CTTCTCTGGATCCTCCCCACCCCCTCCTGCTCCTGGGGCCGGCCCTTTTCTCAGAGATCA 1325
Db      1135 CCTCTAGGTCCCTGACTCCATCTCTGCTGCTCCCAGGGCAGGCCTTTTTCTCAGAGGTCA 1194

Qy      1326 CTCAATAAACCTAAGAACCCTCATAAAAAAAAAAAAAAAA 1365
Db      1195 CTTAATAAACCTAAAATCCTCAAAAAAAAAAAAAAAAAAA 1234
```

```
RESULT 12
BC023068
LOCUS       BC023068                1271 bp    mRNA    linear   ROD 07-AUG-2002
DEFINITION  Mus musculus, Similar to DKFZP586B0621 protein, clone MGC:38635
            IMAGE:5355789, mRNA, complete cds.
ACCESSION   BC023068
VERSION     BC023068.1  GI:18605993
KEYWORDS    MGC.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 1271)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (04-FEB-2002) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Jeffrey Green M.D.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Baylor College of Medicine Human Genome
            Sequencing Center
            Center code: BCM-HGSC
            Web site: http://www.hgsc.bcm.tmc.edu/cdna/
            Contact: amg@bcm.tmc.edu
            Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Hale, S.M.,
            Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M.,
            Richards, S., Gibbs, R.A.

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAK Plate: 54 Row: b Column: 4
            This clone was selected for full length sequencing because it
            passed the following selection criteria: Hexamer frequency ORF
            analysis.
FEATURES             Location/Qualifiers
     source          1. .1271
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /map="FVB/N"
                     /clone="MGC:38635 IMAGE:5355789"
                     /tissue_type="Mammary tumor. C3(1)-Tag model. Infiltrating
                     ductal carcinoma. 5 month old virgin mouse."
                     /clone_lib="NCI_CGAP_Mam6"
                     /lab_host="DH10B"
                     /note="Vector: pCMV-SPORT6"
     CDS             153. .884
                     /codon_start=1
```

```
                 /product="Similar to DKFZP586B0621 protein"
                 /protein_id="AAH23068.1"
                 /db_xref="GI:18605994"
                 /translation="MRPLLALLLLGLVSGSPPLDDNKIPSLCPGQPGLPGTPGHHGSQ
                 GLPGRDGRDGRDGAPGAPGEKGEGGRPGLPGPRGEPGPRGEAGPMGAIGPAGECSVPP
                 RSAFSAKRSESRVPPPADTPLPFDRVLLNEQGHFDPTTGKFTCQVPGVYYFAVHATVY
                 RASLQFDLVKNGQSIASFFQYFGGWPKPASLSGGAMVRLEPEDQVWVQVGVGDYIGIY
                 ASIKTDSTFSGFLVYSDWHSSPVFA"
BASE COUNT       258 a    375 c    378 g    260 t
ORIGIN

  Query Match           50.2%;  Score 691.2;  DB 10;  Length 1271;
  Best Local Similarity  77.9%;  Pred. No. 1.8e-109;
  Matches  916;  Conservative    0;  Mismatches  238;  Indels   22;  Gaps    6;

Qy       189 GAGAGGACCCCGGCGTCCGGGCTCCCGGTGCCAGCGCTATGAGGCCACTCCTCGTCCTGC 248
Db       115 GAGAGGTTCCTGGAGTCTGAGCCTCCGGGGTCACCACCATGAGGCCACTTCTTGCCCTTC 174

Qy       249 TGCTCCTGGGCCTGGCGGCCGGCTCGCCCCCACTGGACGACAACAAGATCCCCAGCCTCT 308
Db       175 TGCTTCTGGGTCTGGTGTCAGGCTCTCCTCCTCTGGACGACAACAAGATCCCCAGCCTGT 234

Qy       309 GCCCGGGGCACCCCGGCCTTCCAGGCACGCCGGGCCACCATGGCAGCCAGGGCTTGCCGG 368
Db       235 GTCCCGGGCAGCCCGGCCTTCCAGGCACACCAGGTCACCATGGCAGCCAAGGCCTGCCTG 294

Qy       369 GCCGCGATGGCCGCGACGGCCGCGACGGCGCGCCCGGGGCTCCGGGAGAGAAAGGCGAGG 428
Db       295 GCCGTGACGGCCGTGATGGCCGCGACGGTGCACCCGGAGCTCCGGGAGAGAAAGGCGAGG 354

Qy       429 GCGGGAGGCCGGGACTGCCGGGACCTCGAGGGGACCCCGGGCCGCGAGGAGAGGCGGGAC 488
Db       355 GCGGGAGACCGGGACTACCTGGGCCACGTGGGGAGCCCGGGCCGCGTGGAGAGGCAGGGC 414

Qy       489 CCGCGGGGCCCACCGGGCCTGCCGGGGAGTGCTCGGTGCCTCCGCGATCCGCCTTCAGCG 548
Db       415 CCATGGGGGCTATCGGGCCTGCGGGGGAGTGCTCGGTACCCCCACGATCAGCCTTCAGTG 474

Qy       549 CCAAGCGCTCCGAGAGCCGGGTGCCTCCGCCGTCTGACGCACCCTTGCCCTTCGACCGCG 608
Db       475 CCAAGCGATCCGAGAGCCGGGTACCTCCGCCAGCCGACACACCCCTACCTTTCGACCGTG 534

Qy       609 TGCTGGTGAACGAGCAGGGACATTACGACGCCGTCACCGGCAAGTTCACCTGCCAGGTGC 668
Db       535 TGCTGCTAAATGAGCAGGGCCATTTCGACCCCACTACTGGCAAGTTCACCTGCCAAGTGC 594

Qy       669 CTGGGGTCTACTACTTCGCCGTCCATGCCACCGTCTACCGGGCCAGCCTGCAGTTTGATC 728
Db       595 CTGGCGTCTACTACTTTGCTGTGCACGCCACTGTCTACCGGGCCAGCTTGCAGTTTGATC 654

Qy       729 TGGTGAAGAATGGCGAATCCATTGCCTCTTTCTTCCAGTTTTTCGGGGGGTGGCCCAAGC 788
Db       655 TTGTCAAAAACGGGCAGTCCATCGCCTCTTTCTTCCAGTATTTTGGGGGGTGGCCCAAGC 714

Qy       789 CAGCCTCGCTCTCGGGGGGGGCCATGGTGAGGCTGGAGCCTGAGGACCAAGTGTGGGTGC 848
Db       715 CAGCCTCGCTCTCAGGGGGTGCGATGGTAAGGCTAGAACCTGAGGACCAGGTGTGGGTGC 774

Qy       849 AGGTGGGTGTGGGTGACTACATTGGCATCTATGCCAGCATCAAGACAGACAGCACCTTCT 908
Db       775 AGGTGGGCGTGGGTGATTACATTGGCATCTATGCCAGCATCAAGACAGACAGTACCTTCT 834

Qy       909 CCGGATTTCTGGTGTACTCCGACTGGCACAGCTCCCCAGTCTTTGCTTAGTGCCCACTGC 968
Db       835 CTGGATTTCTCGTCTATTCTGACTGGCACAGCTCCCCAGTCTTCGCTTAAAACACAGTGA 894

Qy       969 AAAGTGAGCTCATGCTCTCACTCCTAGAAGGAGGGTGTGAGGCTGACAACCAGGTCATCC 1028
Db       895 ACCCGGAGCTGGCACTTGCTCCTCAGTGGAGGGTGTGACACTAACCCGCGCAGCGCATAC 954

Qy      1029 AGGAGGGCTGGCCCCCCTGGAATATTGTGAATGACTAGGGAGGTGGGGTAGAGCACTCTC 1088

Db       955 CAGGAGGGCTGGCCCCCTGGAATATTGTGAATGACTTAGGAAGAGAGGGAG-CCACTTCC 1013

Qy      1089 CGTCCTGCTGCTGGCAAGGAATGGGAACAGTGGCTGTCTGCGATCAGGTCTGGCAGCATG 1148
Db       1014 GGTCCCACTGCTGGCAATGAATGGAGACA--GGCTGTCTGAGGTCA----AGACAGCGTG 1067

Qy      1149 GGGCAGTGGCTGGATTTCTGCCCAAGACCAGAGGAGTGTGCTGTGCTGGCAAGTGTAAGT 1208
Db       1068 GAGCAGTGGCTGGGTTTCTGCCCAGGACTTTA-GAATGCAGTAGGCTGGCAGCTGTGGGT 1126

Qy      1209 CCCCCAGTTGCTCTGGTCCAGGAGCCCACGGTGGGGTGCTCTCTTCCTGGTCCTCTG--- 1265
Db       1127 C-----------CTGGCCCAGGACTCCAAGGTGGGATGCTCCATTCCTAGTCCTGTGCCC 1175

Qy      1266 CTTCTCTGGATCCTCCCCACCCCCTCCTGCTCCTGGGGCCGGCCCTTTTCTCAGAGATCA 1325
Db       1176 CCTCTAGGTCCCTGACTCCATCTCTGCTGCTCCCAGGGCAGGCCTTTTTCTCAGAGGTCA 1235

Qy      1326 CTCAATAAACCTAAGAACCCTCATAAAAAAAAAAAA 1361
Db       1236 CTTAATAAACCTAAAATCCTCAAAAAAAAAAAAAA 1271
```

```
RESULT 13
AP001156
LOCUS       AP001156              198902 bp    DNA     linear   HTG 30-MAY-2000
DEFINITION  Homo sapiens chromosome 11 clone RP11-657A24 map 11q23, WORKING
            DRAFT SEQUENCE, 42 unordered pieces.
ACCESSION   AP001156
VERSION     AP001156.2  GI:8118724
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      Homo sapiens DNA, clone:RP11-657A24.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 198902)
  AUTHORS   Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
            Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
  TITLE     Homo sapiens 198,902 genomic DNA of 11q23
  JOURNAL   Published Only in DataBase (2000)
REFERENCE   2  (bases 1 to 198902)
  AUTHORS   Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
            Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-FEB-2000) Masahira Hattori, The Institute of Physical
            and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
            Kitasato Univ., 1-15-1 Kitasato, Sagamihara, Kanagawa 228-8555,
            Japan (E-mail:hattori@gsc.riken.go.jp,
            URL:http://hgp.gsc.riken.go.jp/, Tel:81-42-778-9923,
            Fax:81-42-778-9924)
COMMENT     On May 31, 2000 this sequence version replaced gi:6997842.
            ---------- Genome Center
                Center: RIKEN Genomic Sciences Center(GSC)
                Center code: RIKEN
                Web site: http://hgp.gsc.riken.go.jp/
                Contact: hattori@gsc.riken.go.jp
            ---------- Project Information
                Center project name: HumDraft11
                Center clone name: RP11-657A24
            ---------- Summary Statistics
                Sequencing vector: PCR products; 100% of reads
                Chemistry: Dye-terminator ET-amersham; 100% of reads
                Assembly program: Phrap; version 0.990329
                Consensus quality: 171788 bases at least Q40
                Consensus quality: 184443 bases at least Q30
                Consensus quality: 191295 bases at least Q20
                Insert size: 194802; sum-of-contigs
                Quality coverage: 4.12x in Q20 bases; sum-of-contigs
            ----------
            NOTE: This is a 'working draft' sequence. It currently consists of
            42 contigs. The true order of the pieces is not known and their
            order in this sequence record is arbitrary. Gaps between the
            contigs are represented as runs N, but the exact sizes of the gaps
```

are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

```
         1    20627 contig of  20627 bp in length
     20728    31817 contig of  11090 bp in length
     31918    41131 contig of   9214 bp in length
     41232    49745 contig of   8514 bp in length
     49846    59457 contig of   9612 bp in length
     59558    66804 contig of   7247 bp in length
     66905    73970 contig of   7066 bp in length
     74071    81007 contig of   6937 bp in length
     81108    87320 contig of   6213 bp in length
     87421    93250 contig of   5830 bp in length
     93351    97965 contig of   4615 bp in length
     98066   102941 contig of   4876 bp in length
    103042   109017 contig of   5976 bp in length
    109118   113391 contig of   4274 bp in length
    113492   118783 contig of   5292 bp in length
    118884   124241 contig of   5358 bp in length
    124342   129766 contig of   5425 bp in length
    129867   135473 contig of   5607 bp in length
    135574   140585 contig of   5012 bp in length
    140686   144027 contig of   3342 bp in length
    144128   147923 contig of   3796 bp in length
    148024   152461 contig of   4438 bp in length
    152562   155444 contig of   2883 bp in length
    155545   158721 contig of   3177 bp in length
    158822   162903 contig of   4082 bp in length
    163004   166121 contig of   3118 bp in length
    166222   169851 contig of   3630 bp in length
    169952   172048 contig of   2097 bp in length
    172149   175332 contig of   3184 bp in length
    175433   177914 contig of   2482 bp in length
    178015   179875 contig of   1861 bp in length
    179976   181604 contig of   1629 bp in length
    181705   183951 contig of   2247 bp in length
    184052   185927 contig of   1876 bp in length
    186028   187870 contig of   1843 bp in length
    187971   189591 contig of   1621 bp in length
    189692   191169 contig of   1478 bp in length
    191270   192527 contig of   1258 bp in length
    192628   194288 contig of   1661 bp in length
    194389   195658 contig of   1270 bp in length
    195759   197707 contig of   1949 bp in length
    197808   198902 contig of   1095 bp in length
Sequence updated (26-May-2000).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 42 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*         1    20627: contig of 20627 bp in length
*     20628 20727: gap of      100 bp
*     20728    31817: contig of 11090 bp in length
*     31818 31917: gap of      100 bp
*     31918    41131: contig of 9214 bp in length
*     41132 41231: gap of      100 bp
*     41232    49745: contig of 8514 bp in length
*     49746 49845: gap of      100 bp
*     49846    59457: contig of 9612 bp in length
*     59458 59557: gap of      100 bp
*     59558    66804: contig of 7247 bp in length
*     66805 66904: gap of      100 bp
*     66905    73970: contig of 7066 bp in length
*     73971 74070: gap of      100 bp
*     74071    81007: contig of 6937 bp in length
*     81008 81107: gap of      100 bp
*     81108    87320: contig of 6213 bp in length
*     87321 87420: gap of      100 bp
*     87421    93250: contig of 5830 bp in length
*     93251 93350: gap of      100 bp
*     93351    97965: contig of 4615 bp in length
*     97966 98065: gap of      100 bp
*     98066   102941: contig of 4876 bp in length
*    102942 103041: gap of      100 bp
*    103042   109017: contig of 5976 bp in length
*    109018 109117: gap of      100 bp
*    109118   113391: contig of 4274 bp in length
*    113392 113491: gap of      100 bp
*    113492   118783: contig of 5292 bp in length
*    118784 118883: gap of      100 bp
*    118884   124241: contig of 5358 bp in length
*    124242 124341: gap of      100 bp
*    124342   129766: contig of 5425 bp in length
*    129767 129866: gap of      100 bp
*    129867   135473: contig of 5607 bp in length
*    135474 135573: gap of      100 bp
*    135574   140585: contig of 5012 bp in length
*    140586 140685: gap of      100 bp
*    140686   144027: contig of 3342 bp in length
*    144028 144127: gap of      100 bp
*    144128   147923: contig of 3796 bp in length
*    147924 148023: gap of      100 bp
*    148024   152461: contig of 4438 bp in length
*    152462 152561: gap of      100 bp
*    152562   155444: contig of 2883 bp in length
*    155445 155544: gap of      100 bp
*    155545   158721: contig of 3177 bp in length
*    158722 158821: gap of      100 bp
*    158822   162903: contig of 4082 bp in length
*    162904 163003: gap of      100 bp
*    163004   166121: contig of 3118 bp in length
*    166122 166221: gap of      100 bp
*    166222   169851: contig of 3630 bp in length
*    169852 169951: gap of      100 bp
*    169952   172048: contig of 2097 bp in length
*    172049 172148: gap of      100 bp
*    172149   175332: contig of 3184 bp in length
*    175333 175432: gap of      100 bp
*    175433   177914: contig of 2482 bp in length
*    177915 178014: gap of      100 bp
*    178015   179875: contig of 1861 bp in length
*    179876 179975: gap of      100 bp
*    179976   181604: contig of 1629 bp in length
*    181605 181704: gap of      100 bp
*    181705   183951: contig of 2247 bp in length
*    183952 184051: gap of      100 bp
*    184052   185927: contig of 1876 bp in length
*    185928 186027: gap of      100 bp
*    186028   187870: contig of 1843 bp in length
*    187871 187970: gap of      100 bp
*    187971   189591: contig of 1621 bp in length
*    189592 189691: gap of      100 bp
*    189692   191169: contig of 1478 bp in length
*    191170 191269: gap of      100 bp
*    191270   192527: contig of 1258 bp in length
*    192528 192627: gap of      100 bp
*    192628   194288: contig of 1661 bp in length
*    194289 194388: gap of      100 bp
*    194389   195658: contig of 1270 bp in length
*    195659 195758: gap of      100 bp
*    195759   197707: contig of 1949 bp in length
*    197708 197807: gap of      100 bp
*    197808   198902: contig of 1095 bp in length.
FEATURES             Location/Qualifiers
     source          1. .198902
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="11"
                     /map="11q23"
                     /clone="RP11-657A24"
     misc_feature    1. .20627
                     /note="assembly_fragment"
```

```
     misc_feature    20728. .31817
                     /note="assembly_fragment"
     misc_feature    31918. .41131
                     /note="assembly_fragment clone_end:SP6 vector_side:right"
     misc_feature    41232. .49745
                     /note="assembly_fragment"

  Query Match             44.5%;  Score 612.8;  DB 2;  Length 198902;
  Best Local Similarity   99.7%;  Pred. No. 3.8e-96;
  Matches  614;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy        735 AGAATGGCGAATCCATTGCCTCTTTCTTCCAGTTTTTCGGGGGGTGGCCCAAGCCAGCCT 794
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     194389 AGAATGGCGAATCCATTGCCTCTTTCTTCCAGTTTTTCGGGGGGTGGCCCAAGCCAGCCT 194448

Qy        795 CGCTCTCGGGGGGGGCCATGGTGAGGCTGGAGCCTGAGGACCAAGTGTGGGTGCAGGTGG 854
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     194449 CGCTCTCGGGGGGGGCCATGGTGAGGCTGGAGCCTGAGGACCAAGTGTGGGTGCAGGTGG 194508

Qy        855 GTGTGGGTGACTACATTGGCATCTATGCCAGCATCAAGACAGACAGCACCTTCTCCGGAT 914
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     194509 GTGTGGGTGACTACATTGGCATCTATGCCAGCATCAAGACAGACAGCACCTTCTCCGGAT 194568

Qy        915 TTCTGGTGTACTCCGACTGGCACAGCTCCCCAGTCTTTGCTTAGTGCCCACTGCAAAGTG 974
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     194569 TTCTGGTGTACTCCGACTGGCACAGCTCCCCAGTCTTTGCTTAGTGCCCACTGCAAAGTG 194628

Qy        975 AGCTCATGCTCTCACTCCTAGAAGGAGGGTGTGAGGCTGACAACCAGGTCATCCAGGAGG 1034
              |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db     194629 AGCTCATGCTCTCACTCCTAGAAGGAGGGTGTGAGGCTGACAACCTGGTCATCCAGGAGG 194688

Qy       1035 GCTGGCCCCCCCTGGAATATTGTGAATGACTAGGGAGGTGGGGTAGAGCACTCTCCGTCCT 1094
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     194689 GCTGGCCCCCCCTGGAATATTGTGAATGACTAGGGAGGTGGGGTAGAGCACTCTCCGTCCT 194748

Qy       1095 GCTGCTGGCAAGGAATGGGAACAGTGGCTGTCTGCGATCAGGTCTGGCAGCATGGGGCAG 1154
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     194749 GCTGCTGGCAAGGAATGGGAACAGTGGCTGTCTGCGATCAGGTCTGGCAGCATGGGGCAG 194808

Qy       1155 TGGCTGGATTTCTGCCCAAGACCAGAGGAGTGTGCTGTGCTGGCAAGTGTAAGTCCCCCA 1214
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     194809 TGGCTGGATTTCTGCCCAAGACCAGAGGAGTGTGCTGTGCTGGCAAGTGTAAGTCCCCCA 194868

Qy       1215 GTTGCTCTGGTCCAGGAGCCCACGGTGGGGTGCTCTCTTCCTGGTCCTCTGCTTCTCTGG 1274
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     194869 GTTGCTCTGGTCCAGGAGCCCACGGTGGGGTGCTCTCTTCCTGGTCCTCTGCTTCTCTGG 194928

Qy       1275 ATCCTCCCCACCCCCTCCTGCTCCTGGGGCCGGCCCTTTTCTCAGAGATCACTCAATAAA 1334
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     194929 ATCCTCCCCACCCCCTCCTGCTCCTGGGGCCGGCCCTTTTCTCAGAGATCACTCAATAAA 194988

Qy       1335 CCTAAGAACCCTCATA 1350
              |||||||||||||| ||
Db     194989 CCTAAGAACCCTCCTA 195004


RESULT 14
AC112557
LOCUS       AC112557               173657 bp    DNA     linear   HTG 14-JUL-2002
DEFINITION  Rattus norvegicus clone CH230-904, *** SEQUENCING IN PROGRESS ***,
            72 unordered pieces.
ACCESSION   AC112557
VERSION     AC112557.2  GI:21735369
KEYWORDS    HTG; HTGS_PHASE1.
SOURCE      Norway rat.
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 173657)
  AUTHORS   Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
            Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Ayele,M., Banks,T.,
            Barbaria,J., Benton,J., Bimage,K., Blankenburg,K., Bonnin,D.,
            Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,
            Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
            Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
            Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
            Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
            Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
            Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
            Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
            Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
            Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
            Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
            Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
            Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,
            Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B.,
            Homsi,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E.,
            Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,
            Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,
            Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
            Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Loulseged,H.,
            Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
            Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,
            Massey,E., Mawhiney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M.,
            Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,
            Moser,M., Neal,D., Newtson,J., Newtson,N., Nguyen,A., Nguyen,N.,
            Nguyen,N., Nickerson,E., Nwokenkwo,S., Oguh,M., Okwuonu,G.,
            Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
            Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,
            Rives,M., Rojas,A., Rojubokan,I., Rolfe,M., Ruiz,S., Savery,G.,
            Scherer,S., Scott,G., Shen,H., Shooshtari,N., Sisson,I.,
            Sodergren,E., Sonaike,T., Sparks,A., Stanley,H., Stone,H.,
            Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,
            Tansey,J., Taylor,C., Taylor,T., Telfrod,B., Thomas,N., Thomas,S.,
            Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q.,
            Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,
            Williams,G., Williamson,A., Wleczyk,R., Wooden,S., Worley,K.,
            Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
            Weinstock,G. and Gibbs,R.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 173657)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (22-FEB-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE   3  (bases 1 to 173657)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (14-JUL-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
COMMENT     On Jul 12, 2002 this sequence version replaced gi:18860143.
            -------------- Genome Center
                Center: Baylor College of Medicine
                Center code: BCM
                Web site: http://www.hgsc.bcm.tmc.edu/
                Contact: hgsc-help@bcm.tmc.edu
            -------------- Project Information
                Center project name: GKAB
                Center clone name: CH230-904
            -------------- Summary Statistics
                Sequencing vector: Plasmid;
                Chemistry: Dye-terminator Big Dye: 100% of reads
                Assembly program: Phrap; version 0.990329
                Consensus quality: 105614 bases at least Q40
                Consensus quality: 111210 bases at least Q30
                Consensus quality: 115587 bases at least Q20
            --------------
            * NOTE: Estimated insert size may differ from sequence length
            *   (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 72 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
```

```
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*        1    1137: contig of 1137 bp in length
*     1138    1237: gap of unknown length
*     1238    2720: contig of 1483 bp in length
*     2721    2820: gap of unknown length
*     2821    3984: contig of 1164 bp in length
*     3985    4084: gap of unknown length
*     4085    5158: contig of 1074 bp in length
*     5159    5258: gap of unknown length
*     5259    6306: contig of 1048 bp in length
*     6307    6406: gap of unknown length
*     6407    7523: contig of 1117 bp in length
*     7524    7623: gap of unknown length
*     7624    8631: contig of 1008 bp in length
*     8632    8731: gap of unknown length
*     8732    9806: contig of 1075 bp in length
*     9807    9906: gap of unknown length
*     9907   11155: contig of 1249 bp in length
*    11156   11255: gap of unknown length
*    11256   12568: contig of 1313 bp in length
*    12569   12668: gap of unknown length
*    12669   13704: contig of 1036 bp in length
*    13705   13804: gap of unknown length
*    13805   14931: contig of 1127 bp in length
*    14932   15031: gap of unknown length
*    15032   16084: contig of 1053 bp in length
*    16085   16184: gap of unknown length
*    16185   17858: contig of 1674 bp in length
*    17859   17958: gap of unknown length
*    17959   19035: contig of 1077 bp in length
*    19036   19135: gap of unknown length
*    19136   20291: contig of 1156 bp in length
*    20292   20391: gap of unknown length
*    20392   21643: contig of 1252 bp in length
*    21644   21743: gap of unknown length
*    21744   22846: contig of 1103 bp in length
*    22847   22946: gap of unknown length
*    22947   24684: contig of 1738 bp in length
*    24685   24784: gap of unknown length
*    24785   26405: contig of 1621 bp in length
*    26406   26505: gap of unknown length
*    26506   27716: contig of 1211 bp in length
*    27717   27816: gap of unknown length
*    27817   28876: contig of 1060 bp in length
*    28877   28976: gap of unknown length
*    28977   30019: contig of 1043 bp in length
*    30020   30119: gap of unknown length
*    30120   31550: contig of 1431 bp in length
*    31551   31650: gap of unknown length
*    31651   33128: contig of 1478 bp in length
*    33129   33228: gap of unknown length
*    33229   34763: contig of 1535 bp in length
*    34764   34863: gap of unknown length
*    34864   36404: contig of 1541 bp in length
*    36405   36504: gap of unknown length
*    36505   37957: contig of 1453 bp in length
*    37958   38057: gap of unknown length
*    38058   40293: contig of 2236 bp in length
*    40294   40393: gap of unknown length
*    40394   41698: contig of 1305 bp in length
*    41699   41798: gap of unknown length
*    41799   43155: contig of 1357 bp in length
*    43156   43255: gap of unknown length
*    43256   45193: contig of 1938 bp in length
*    45194   45293: gap of unknown length
*    45294   47217: contig of 1924 bp in length
*    47218   47317: gap of unknown length
*    47318   49231: contig of 1914 bp in length
*    49232   49331: gap of unknown length
*    49332   51205: contig of 1874 bp in length
*    51206   51305: gap of unknown length
*    51306   53117: contig of 1812 bp in length
*    53118   53217: gap of unknown length
*    53218   55542: contig of 2325 bp in length
*    55543   55642: gap of unknown length
*    55643   57451: contig of 1809 bp in length
*    57452   57551: gap of unknown length
*    57552   60851: contig of 3300 bp in length
*    60852   60951: gap of unknown length
*    60952   62718: contig of 1767 bp in length
*    62719   62818: gap of unknown length
*    62819   64976: contig of 2158 bp in length
*    64977   65076: gap of unknown length
*    65077   67855: contig of 2779 bp in length
*    67856   67955: gap of unknown length
*    67956   71313: contig of 3358 bp in length
*    71314   71413: gap of unknown length
*    71414   73659: contig of 2246 bp in length
*    73660   73759: gap of unknown length
*    73760   75549: contig of 1790 bp in length
*    75550   75649: gap of unknown length
*    75650   79106: contig of 3457 bp in length
*    79107   79206: gap of unknown length
*    79207   83551: contig of 4345 bp in length
*    83552   83651: gap of unknown length
*    83652   86053: contig of 2402 bp in length
*    86054   86153: gap of unknown length
*    86154   88659: contig of 2506 bp in length
*    88660   88759: gap of unknown length
*    88760   90078: contig of 1319 bp in length
*    90079   90178: gap of unknown length
*    90179   92201: contig of 2023 bp in length
*    92202   92301: gap of unknown length
*    92302   94877: contig of 2576 bp in length
*    94878   94977: gap of unknown length
*    94978   96946: contig of 1969 bp in length
*    96947   97046: gap of unknown length
*    97047   99897: contig of 2851 bp in length

Query Match             37.3%;  Score 513.4;  DB 2;  Length 173657;
Best Local Similarity   78.9%;  Pred. No. 5.2e-79;
Matches  717;  Conservative    0;  Mismatches  171;  Indels   21;  Gaps    8;

Qy        438 CGGGACTGCCGGGACCTCGAGGGGACCCCGGGCCGCGAGGAGAGGCGGGACCCGCGGGGC 497
              | ||||| || || || || ||||| |||||||||| |||||||| ||||| | ||||
Db      43765 CAGGACTACCTGGGCCACGTGGGGAGCCCGGGCCGCGTGGAGAGGCAGGACCTGTGGGGG 43824

Qy        498 CCACCGGGCCTGCCGGGGAGTGCTCGGTGCCTCCGCGATCCGCCTTCAGCGCCAAGCGCT 557
              | | ||||||||| ||||||||||||||||| || ||||| |||||||| |||||||| |
Db      43825 CTATCGGGCCTGCGGGGGAGTGCTCGGTGCCCCCACGATCAGCCTTCAGTGCCAAGCGAT 43884

Qy        558 CCGAGAGCCGGGTGCCTCCGCCGTCTGACGCACCCTTGCCCTTCGACCGCGTGCTGGTGA 617
              | |||||||||| |||||||| | ||| ||||| | |||||||||| |||||| | |
Db      43885 CAGAGAGCCGGGTACCTCCGCCAGCCGACACACCCCTACCCTTCGACCGTGTGCTGCTCA 43944

Qy        618 ACGAGCAGGGACATTACGACGCCGTCACCGGCAAGTTCACCTGCCAGGTGCCTGGGGTCT 677
              | |||||||||||||||| ||| |||||||||||||||||||| |||||||| ||||
Db      43945 ATGAGCAGGGACATTACGATGCCACTACCGGCAAGTTCACCTGCCAAGTGCCTGGTGTCT 44004

Qy        678 ACTACTTCGCCGTCCATGCCACCGTCTACCGGGCCAGCCTGCAGTTTGATCTGGTGAAGA 737
              ||||||| || |||||||||| |||||||||||||||| |||||||||| || || |
Db      44005 ACTACTTTGCTGTCCATGCCACTGTCTACCGGGCCAGCCTACAGTTTGATCTTGTCAAAA 44064

Qy        738 ATGGCGAATCCATTGCCTCTTTCTTCCAGTTTTTCGGGGGGTGGCCCAAGCCAGCCTCGC 797
              ||||| ||||||| || |||||||||||||||| |||||||||| ||||||||||||
Db      44065 ATGGCCAATCCATAGCTTCTTTCTTCCAGTTTTTGGGGGGTGGCCAAAGCCAGCCTCGC 44124

Qy        798 TCTCGGGGGGGGCCATGGTGAGGCTGGAGCCTGAGGACCAAGTGTGGGTGCAGGTGGGTG 857
              |||| ||||| || |||||||||| || |||||||||| || ||||| ||||||||||
Db      44125 TCTCAGGGGGTGCGATGGTGAGGCTAGAACCTGAGGACCAGGTATGGGTTCAGGTGGGTG 44184

Qy        858 TGGGTGACTACATTGGCATCTATGCCAGCATCAAGACAGACAGCACCTTCTCCGGATTTC 917
```

```
Db      44185 TGGGTGATTACATTGGCATCTATGCCAGCATCAAAACAGACAGTACCTTCTCTGGATTTC 44244

Qy        918 TGGTGTACTCCGACTGGCACAGCTCCCCAGTCTTTGCTTAGTGCCCACTGCAAAGTGAGC 977

Db      44245 TCGTCTATTCTGACTGGCACAGCTCCCCAGTCTTCGCTTAAAATACAGTGAACCCGGAGC 44304

Qy        978 TCATGCTCTCACTCCTAGAAGGAGGGTGTGAGGCTGACAACCAGGTCATCCAGGAGGGCT 1037

Db      44305 T--GGCACTTGCTCCTAG-TGGAGGGTGTGACACTGACCAGCGCGCATACCAGGAGGGCT 44361

Qy       1038 GGCCCCCCTGGAATATTGTGAATGACTAGGGAGGTGGGGTAGAGCACTCTCCGTCCTGCT 1097

Db      44362 GG-CCCCCTGGACTATTGTGAATGACTAAGAAGAGAGGG--AGCCACCTCTGGTCCCCCT 44418

Qy       1098 GCTGGCAAGGAATGGGAACAGTGGCTGTCTGCGATCAGGTCTGGCAGCATGGGGCAGTGG 1157

Db      44419 CCTGGCAATGAATAG--TCAGAGGCTGTCTGAGGTCAGGACAGACAGCGTGGAGCAGTTG 44476

Qy       1158 CTGGATTTCTGCCCAAGACCAGAGGAGTGTGCTGTGCTGGCAAGTGTAAGTCCCCCAGTT 1217

Db      44477 CTGGGTTTCTGCCCAGGACTTTAGAATGCAGCAG-GCTGGCAGCTGT----------GGA 44525

Qy       1218 GCTCTGGTCCAGGAGCCCACGGTGGGGTGCTCTCTTCCTGGTCCTCTGCTTCTCTGGATC 1277

Db      44526 ATCCTGGCTCAGGACTCCAAGATGGGATGCTCCCTTCCTAGTCCTGTGCCCCTCTAGGTC 44585

Qy       1278 CTCCCCACC--CCCTCCTGCTCCTGGGGCCGGCCCTTTTCTCAGAGATCACTCAATAAAC 1335

Db      44586 TTTGACTCCATCTCTCCTGCTCCCAAGGCAGGCCTTTTTCTCATAGGTCACTTAATAAAC 44645

Qy       1336 CTAAGAACC 1344

Db      44646 CTACAATCC 44654
```

```
RESULT 15
AR138195
LOCUS       AR138195                729 bp    DNA     linear   PAT 16-JUN-2001
DEFINITION  Sequence 10 from patent US 6197930.
ACCESSION   AR138195
VERSION     AR138195.1  GI:14479704
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 729)
  AUTHORS   Sheppard,P.O. and Humes,J.M.
  TITLE     Adipocyte-specific protein homologs
  JOURNAL   Patent: US 6197930-A 10 06-MAR-2001;
FEATURES             Location/Qualifiers
     source          1. .729
                     /organism="unknown"
BASE COUNT       77 a    101 c    160 g     81 t    310 others
ORIGIN
```

Query Match 36.6%; Score 504; DB 6; Length 729;
Best Local Similarity 57.6%; Pred. No. 3.4e-77;
Matches 419; Conservative 145; Mismatches 164; Indels 0; Gaps 0;

```
Qy        227 ATGAGGCCACTCCTCGTCCTGCTGCTCCTGGGCCTGGCGGCCGGCTCGCCCCCACTGGAC 286

Db          1 ATGMGNCCNYTNYTNGTNYTNYTNYTNYTNGGNYTNGCNGCNGGNWSNCCNCCNYTNGAY 60

Qy        287 GACAACAAGATCCCCAGCCTCTGCCCGGGGCACCCCGGCCTTCCAGGCACGCCGGGCCAC 346

Db         61 GAYAAYAARATHCCNWSNYTNTGYCCNGGNCAYCCNGGNYTNCCNGGNACNCCNGGNCAY 120

Qy        347 CATGGCAGCCAGGGCTTGCCGGGCCGCGATGGCCGCGACGGCCGCGACGGCGCGCCCGGG 406

Db        121 CAYGGNWSNCARGGNYTNCCNGGNMGNGAYGGNMGNGAYGGNMGNGAYGGNGCNCCNGGN 180

Qy        407 GCTCCGGGAGAGAAAGGCGAGGGCGGGAGGCCGGGACTGCCGGGACCTCGAGGGGACCCC 466

Db        181 GCNCCNGGNGARAARGGNGARGGNGGNMGNCCNGGNYTNCCNGGNCCNMGNGGNGAYCCN 240

Qy        467 GGGCCGCGAGGAGAGGCGGGACCCGCGGGGCCCACCGGGCCTGCCGGGGAGTGCTCGGTG 526

Db        241 GGNCCNMGNGGNGARGCNGGNCCNGCNGGNCCNACNGGNCCNGCNGGNGARTGYWSNGTN 300

Qy        527 CCTCCGCGATCCGCCTTCAGCGCCAAGCGCTCCGAGAGCCGGGTGCCTCCGCCGTCTGAC 586

Db        301 CCNCCNMGNWSNGCNTTYWSNGCNAARMGNWSNGARWSNMGNGTNCCNCCNCCNWSNGAY 360

Qy        587 GCACCCTTGCCCTTCGACCGCGTGCTGGTGAACGAGCAGGGACATTACGACGCCGTCACC 646

Db        361 GCNCCNYTNCCNTTYGAYMGNGTNYTNGTNAAYGARCARGGNCAYTAYGAYGCNGTNACN 420

Qy        647 GGCAAGTTCACCTGCCAGGTGCCTGGGGTCTACTACTTCGCCGTCCATGCCACCGTCTAC 706

Db        421 GGNAARTTYACNTGYCARGTNCCNGGNGTNTAYTAYTTYGCNGTNCAYGCNACNGTNTAY 480

Qy        707 CGGGCCAGCCTGCAGTTTGATCTGGTGAAGAATGGCGAATCCATTGCCTCTTTCTTCCAG 766

Db        481 MGNGCNWSNYTNCARTTYGAYYTNGTNAARAAYGGNGARWSNATHGCNWSNTTYTTYCAR 540

Qy        767 TTTTTCGGGGGGTGGCCCAAGCCAGCCTCGCTCTCGGGGGGGGCCATGGTGAGGCTGGAG 826

Db        541 TTYTTYGGNGGNTGGCCNAARCCNGCNWSNYTNWSNGGNGGNGCNATGGTNMGNYTNGAR 600

Qy        827 CCTGAGGACCAAGTGTGGGTGCAGGTGGGTGTGGGTGACTACATTGGCATCTATGCCAGC 886

Db        601 CCNGARGAYCARGTNTGGGTNCARGTNGGNGTNGGNGAYTAYATHGGNATHTAYGCNWSN 660

Qy        887 ATCAAGACAGACAGCACCTTCTCCGGATTTCTGGTGTACTCCGACTGGCACAGCTCCCCA 946

Db        661 ATHAARACNGAYWSNACNTTYWSNGGNTTYYTNGTNTAYWSNGAYTGGCAYWSNWSNCCN 720

Qy        947 GTCTTTGC 954

Db        721 GTNTTYGC 728
```

Search completed: June 22, 2003, 04:43:30
Job time : 3606 secs

GenCore version 5.1.6


OM nucleic - nucleic search, using sw model

Run on: June 22, 2003, 02:42:34 ; Search time 349 Seconds
(without alignments)
8885.392 Million cell updates/sec

Title: US-09-943-851A-41
Perfect score: 1377
Sequence: 1 gactagttctcttggagtct.....aaaaaaaaaaaaaaaaaaa 1377

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 1377 | 100.0 | 1377 | 20 | AAX87258 | cDNA clone encodin |
| 2 | 1377 | 100.0 | 1377 | 20 | AAX80052 | Human PRO344 nucle |
| 3 | 1377 | 100.0 | 1377 | 21 | AAD01241 | Human PRO344 prote |
| 4 | 1377 | 100.0 | 1377 | 21 | AAA46907 | cDNA encoding nove |
| 5 | 1377 | 100.0 | 1377 | 21 | AAA49560 | Human PRO344 cDNA. |
| 6 | 1377 | 100.0 | 1377 | 22 | AAS21424 | Human cDNA sequenc |
| 7 | 1375.4 | 99.9 | 1377 | 21 | AAC58626 | Human PRO344 prote |
| 8 | 1322.2 | 96.0 | 1347 | 20 | AAX24684 | Human adipocyte-sp |
| 9 | 1296.6 | 94.2 | 1338 | 22 | AAF44970 | Human TANGO 253 co |
| 10 | 1295 | 94.0 | 1338 | 22 | AAF44998 | Human secreted pro |
| 11 | 1295 | 94.0 | 1338 | 22 | AAF44999 | Human secreted pro |
| 12 | 1295 | 94.0 | 1338 | 22 | AAF45000 | Human secreted pro |
| 13 | 1295 | 94.0 | 1338 | 22 | AAF45001 | Human secreted pro |
| 14 | 1084.8 | 78.8 | 3248 | 24 | AAL39624 | Human secreted pro |
| 15 | 923.6 | 67.1 | 1082 | 23 | AAS76911 | DNA encoding novel |
| 16 | 763 | 55.4 | 764 | 24 | AAS62228 | cDNA sequence #15 |
| 17 | 717 | 52.1 | 728 | 22 | AAF44971 | Human TANGO 253 OR |
| 18 | 717 | 52.1 | 728 | 22 | AAF44997 | Human secreted pro |
| 19 | 715.4 | 52.0 | 728 | 22 | AAF44994 | Human secreted pro |
| 20 | 715.4 | 52.0 | 728 | 22 | AAF44995 | Human secreted pro |
| 21 | 715.4 | 52.0 | 728 | 22 | AAF44996 | Human secreted pro |
| 22 | 692.8 | 50.3 | 1263 | 22 | AAF44972 | Murine TANGO 253 O |
| 23 | 691.2 | 50.2 | 1263 | 22 | AAF45006 | Murine secreted pr |
| 24 | 691.2 | 50.2 | 1263 | 22 | AAF45007 | Murine secreted pr |
| 25 | 691.2 | 50.2 | 1263 | 22 | AAF45008 | Murine secreted pr |
| 26 | 691.2 | 50.2 | 1263 | 22 | AAF45009 | Murine secreted pr |
| 27 | 606 | 44.0 | 1001 | 21 | AAZ61745 | cDNA encoding rat |
| 28 | 606 | 44.0 | 1001 | 22 | AAC99678 | Skin cell cDNA, SE |
| 29 | 606 | 44.0 | 1001 | 24 | ABL34830 | Rat cDNA isolated |
| 30 | 606 | 44.0 | 1015 | 21 | AAZ61635 | cDNA encoding rat |
| 31 | 606 | 44.0 | 1015 | 22 | AAC99568 | Skin cell cDNA, SE |
| 32 | 606 | 44.0 | 1015 | 24 | ABL34720 | Rat cDNA isolated |
| 33 | 596 | 43.3 | 602 | 22 | AAH33447 | Human colon cancer |
| 34 | 564.2 | 41.0 | 729 | 22 | AAF44973 | Murine TANGO 253 c |
| 35 | 562.6 | 40.9 | 729 | 22 | AAF45038 | Murine secreted pr |
| 36 | 562.6 | 40.9 | 729 | 22 | AAF45039 | Murine secreted pr |
| 37 | 562.6 | 40.9 | 729 | 22 | AAF45040 | Murine secreted pr |
| 38 | 562.6 | 40.9 | 729 | 22 | AAF45041 | Murine secreted pr |
| 39 | 504 | 36.6 | 729 | 20 | AAX24685 | Degenerate DNA enc |
| 40 | 347 | 25.2 | 502 | 21 | AAZ98154 | Human signal pepti |
| 41 | 257.8 | 18.7 | 601 | 22 | AAF44991 | Rat TANGO 253 codi |
| 42 | 247.4 | 18.0 | 536 | 23 | AAS76910 | DNA encoding novel |
| 43 | 149.6 | 10.9 | 393 | 21 | AAZ61631 | cDNA encoding rat |
| 44 | 149.6 | 10.9 | 393 | 22 | AAC99564 | Skin cell cDNA, SE |
| 45 | 149.6 | 10.9 | 393 | 24 | ABL34716 | Rat cDNA isolated |

ALIGNMENTS

RESULT 1
AAX87258
ID AAX87258 standard; cDNA; 1377 BP.
XX
AC AAX87258;
XX
DT 27-SEP-1999 (first entry)
XX
DE cDNA clone encoding human PRO344, amplified in tumour cells.
XX
KW PRO344; UNQ303; cancer; tumour; diagnosis; therapy; human; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 227..958
FT /*tag= a
FT sig_peptide 227..271
FT /*tag= b
FT mat_peptide 272..955
FT /*tag= c
XX
PN WO9935170-A2.
XX
PD 15-JUL-1999.
XX
PF 05-JAN-1999; 99WO-US00106.
XX
PR 20-NOV-1998; 98US-0109304.
PR 05-JAN-1998; 98US-0070440.
PR 29-APR-1998; 98US-0083500.
PR 22-MAY-1998; 98US-0086414.

PR 10-JUN-1998; 98US-0088742.
PR 10-NOV-1998; 98US-0107783.
XX
PA (GETH ) GENENTECH INC.
XX
PI Botstein D, Goddard A, Gurney AL, Hillan KJ, Lawrence DA;
PI Roy MA, Wood WI;
XX
DR WPI; 1999-430385/36.
DR P-PSDB; AAY06481.
XX
PT Antibody against proteins expressed in neoplastic cells, useful for
PT tumor diagnosis and treatment
XX
PS Example 1; Fig 9; 162pp; English.
XX
CC This is the nucleotide sequence of cDNA clone DNA40592 (ATCC 209492)
CC coding for human PRO344 (UNQ303) (see AAY06481). The clone was
CC isolated from a foetal kidney library. Amplification of DNA40592
CC was observed in primary lung tumours and in primary colon tumours,
CC suggesting an association with tumour formation or growth.
CC Antagonists (e.g. antibodies) directed against PRO344 may have
CC utility in cancer therapy. The invention identifies 14 genes (see
CC AAX87254-67) that are amplified in the genome of tumour cells. Such
CC amplification is expected to be associated with overexpression of
CC the gene product and to contribute to tumorigenesis. The encoded
CC proteins (see AAY06477-90) may be useful targets for the diagnosis
CC and/or treatment (including prevention) of certain cancers, and may
CC act as predictors of the prognosis of tumour treatment.
XX
SQ Sequence 1377 BP; 251 A; 422 C; 472 G; 232 T; 0 other;

Query Match 100.0%; Score 1377; DB 20; Length 1377;
Best Local Similarity 100.0%; Pred. No. 1.7e-244;
Matches 1377; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 GACTAGTTCTCTTGGAGTCTGGGAGGAGGAAAGCGGAGCCGGCAGGGAGCGAACCAGGAC 60
Db       1 GACTAGTTCTCTTGGAGTCTGGGAGGAGGAAAGCGGAGCCGGCAGGGAGCGAACCAGGAC 60

Qy      61 TGGGGTGACGGCAGGGCAGGGGGCGCCTGGCCGGGGAGAAGCGCGGGGGCTGGAGCACCA 120
Db      61 TGGGGTGACGGCAGGGCAGGGGGCGCCTGGCCGGGGAGAAGCGCGGGGGCTGGAGCACCA 120

Qy     121 CCAACTGGAGGGTCCGGAGTAGCGAGCGCCCCGAAGGAGGCCATCGGGGAGCCGGGAGGG 180
Db     121 CCAACTGGAGGGTCCGGAGTAGCGAGCGCCCCGAAGGAGGCCATCGGGGAGCCGGGAGGG 180

Qy     181 GGGACTGCGAGAGGACCCCGGCGTCCGGGCTCCCGGTGCCAGCGCTATGAGGCCACTCCT 240
Db     181 GGGACTGCGAGAGGACCCCGGCGTCCGGGCTCCCGGTGCCAGCGCTATGAGGCCACTCCT 240

Qy     241 CGTCCTGCTGCTCCTGGGCCTGGCGGCCGGCTCGCCCCCACTGGACGACAACAAGATCCC 300
Db     241 CGTCCTGCTGCTCCTGGGCCTGGCGGCCGGCTCGCCCCCACTGGACGACAACAAGATCCC 300

Qy     301 CAGCCTCTGCCCGGGGCACCCCGGCCTTCCAGGCACGCCGGGCCACCATGGCAGCCAGGG 360
Db     301 CAGCCTCTGCCCGGGGCACCCCGGCCTTCCAGGCACGCCGGGCCACCATGGCAGCCAGGG 360

Qy     361 CTTGCCGGGCCGCGATGGCCGCGACGGCCGCGACGGCGCGCCCGGGGCTCCGGGAGAGAA 420
Db     361 CTTGCCGGGCCGCGATGGCCGCGACGGCCGCGACGGCGCGCCCGGGGCTCCGGGAGAGAA 420

Qy     421 AGGCGAGGGCGGGAGGCCGGGACTGCCGGGACCTCGAGGGGACCCCGGGCCGCGAGGAGA 480
Db     421 AGGCGAGGGCGGGAGGCCGGGACTGCCGGGACCTCGAGGGGACCCCGGGCCGCGAGGAGA 480

Qy     481 GGCGGGACCCGCGGGGCCCACCGGGCCTGCCGGGGAGTGCTCGGTGCCTCCGCGATCCGC 540
Db     481 GGCGGGACCCGCGGGGCCCACCGGGCCTGCCGGGGAGTGCTCGGTGCCTCCGCGATCCGC 540

Qy     541 CTTCAGCGCCAAGCGCTCCGAGAGCCGGGTGCCTCCGCCGTCTGACGCACCCTTGCCCTT 600
Db     541 CTTCAGCGCCAAGCGCTCCGAGAGCCGGGTGCCTCCGCCGTCTGACGCACCCTTGCCCTT 600

Qy     601 CGACCGCGTGCTGGTGAACGAGCAGGGACATTACGACGCCGTCACCGGCAAGTTCACCTG 660
Db     601 CGACCGCGTGCTGGTGAACGAGCAGGGACATTACGACGCCGTCACCGGCAAGTTCACCTG 660

Qy     661 CCAGGTGCCTGGGGTCTACTACTTCGCCGTCCATGCCACCGTCTACCGGGCCAGCCTGCA 720
Db     661 CCAGGTGCCTGGGGTCTACTACTTCGCCGTCCATGCCACCGTCTACCGGGCCAGCCTGCA 720

Qy     721 GTTTGATCTGGTGAAGAATGGCGAATCCATTGCCTCTTTCTTCCAGTTTTTCGGGGGGTG 780
Db     721 GTTTGATCTGGTGAAGAATGGCGAATCCATTGCCTCTTTCTTCCAGTTTTTCGGGGGGTG 780

Qy     781 GCCCAAGCCAGCCTCGCTCTCGGGGGGGGCCATGGTGAGGCTGGAGCCTGAGGACCAAGT 840
Db     781 GCCCAAGCCAGCCTCGCTCTCGGGGGGGGCCATGGTGAGGCTGGAGCCTGAGGACCAAGT 840

Qy     841 GTGGGTGCAGGTGGGTGTGGGTGACTACATTGGCATCTATGCCAGCATCAAGACAGACAG 900
Db     841 GTGGGTGCAGGTGGGTGTGGGTGACTACATTGGCATCTATGCCAGCATCAAGACAGACAG 900

Qy     901 CACCTTCTCCGGATTTCTGGTGTACTCCGACTGGCACAGCTCCCCAGTCTTTGCTTAGTG 960
Db     901 CACCTTCTCCGGATTTCTGGTGTACTCCGACTGGCACAGCTCCCCAGTCTTTGCTTAGTG 960

Qy     961 CCCACTGCAAAGTGAGCTCATGCTCTCACTCCTAGAAGGAGGGTGTGAGGCTGACAACCA 1020
Db     961 CCCACTGCAAAGTGAGCTCATGCTCTCACTCCTAGAAGGAGGGTGTGAGGCTGACAACCA 1020

Qy    1021 GGTCATCCAGGAGGGCTGGCCCCCCTGGAATATTGTGAATGACTAGGGAGGTGGGGTAGA 1080
Db    1021 GGTCATCCAGGAGGGCTGGCCCCCCTGGAATATTGTGAATGACTAGGGAGGTGGGGTAGA 1080

Qy    1081 GCACTCTCCGTCCTGCTGCTGGCAAGGAATGGGAACAGTGGCTGTCTGCGATCAGGTCTG 1140
Db    1081 GCACTCTCCGTCCTGCTGCTGGCAAGGAATGGGAACAGTGGCTGTCTGCGATCAGGTCTG 1140

Qy    1141 GCAGCATGGGGCAGTGGCTGGATTTCTGCCCAAGACCAGAGGAGTGTGCTGTGCTGGCAA 1200
Db    1141 GCAGCATGGGGCAGTGGCTGGATTTCTGCCCAAGACCAGAGGAGTGTGCTGTGCTGGCAA 1200

Qy    1201 GTGTAAGTCCCCCAGTTGCTCTGGTCCAGGAGCCCACGGTGGGGTGCTCTCTTCCTGGTC 1260
Db    1201 GTGTAAGTCCCCCAGTTGCTCTGGTCCAGGAGCCCACGGTGGGGTGCTCTCTTCCTGGTC 1260

Qy    1261 CTCTGCTTCTCTGGATCCTCCCCACCCCCTCCTGCTCCTGGGGCCGGCCCTTTTCTCAGA 1320
Db    1261 CTCTGCTTCTCTGGATCCTCCCCACCCCCTCCTGCTCCTGGGGCCGGCCCTTTTCTCAGA 1320

Qy    1321 GATCACTCAATAAACCTAAGAACCCTCATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1377
Db    1321 GATCACTCAATAAACCTAAGAACCCTCATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1377
```

RESULT 2
AAX80052
ID AAX80052 standard; cDNA; 1377 BP.
XX
AC AAX80052;
XX
DT 12-AUG-1999 (first entry)
XX
DE Human PRO344 nucleotide sequence.
XX
KW Human; PRO protein; tumour necrosis factor family; TNF; cytokine;
KW secreted protein; transmembrane protein; inflammation disorder; ss.
XX
OS Homo sapiens.
XX
PN WO9928462-A2.
XX

PD 10-JUN-1999.
XX
PF 01-DEC-1998; 98WO-US25108.
XX
PR 25-FEB-1998; 98US-0075945.
PR 03-DEC-1997; 97US-0067411.
PR 11-DEC-1997; 97US-0069278.
PR 11-DEC-1997; 97US-0069334.
PR 11-DEC-1997; 97US-0069335.
PR 12-DEC-1997; 97US-0069425.
PR 16-DEC-1997; 97US-0069694.
PR 16-DEC-1997; 97US-0069696.
PR 16-DEC-1997; 97US-0069702.
PR 17-DEC-1997; 97US-0069870.
PR 17-DEC-1997; 97US-0069873.
PR 18-DEC-1997; 97US-0068017.
PR 05-JAN-1998; 98US-0070440.
PR 09-FEB-1998; 98US-0074086.
PR 09-FEB-1998; 98US-0074092.
XX
PA (GETH ) GENENTECH INC.
XX
PI Baker KP, Chen J, Goddard A, Gurney AL, Wood WI;
PI Yuan J;
XX
DR WPI; 1999-371118/31.
DR P-PSDB; AAY17827.
XX
PT Nucleic acids encoding PRO secreted and transmembrane proteins
XX
PS Claim 2; Fig 20; 123pp; English.
XX
CC The present invention describes nucleic acids encoding PRO secreted and
CC transmembrane proteins used therapeutically. The PRO proteins have
CC cytostatic, anti-inflammatory, anti-proliferative and immunosuppressive
CC activity. The proteins and polynucleotides can be used in therapy,
CC identification of homologues, raising antibodies and design of probes
CC and primers. They can be used in a range of diseases related to proteins
CC that they have homology with, e.g. a PRO protein having homology to
CC complement proteins may be used in inflammatory responses.
XX
SQ Sequence 1377 BP; 251 A; 422 C; 472 G; 232 T; 0 other;

Query Match 100.0%; Score 1377; DB 20; Length 1377;
Best Local Similarity 100.0%; Pred. No. 1.7e-244;
Matches 1377; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 GACTAGTTCTCTTGGAGTCTGGGAGGAGGAAAGCGGAGCCGGCAGGGAGCGAACCAGGAC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GACTAGTTCTCTTGGAGTCTGGGAGGAGGAAAGCGGAGCCGGCAGGGAGCGAACCAGGAC 60

Qy         61 TGGGGTGACGGCAGGGCAGGGGGCGCCTGGCCGGGGAGAAGCGCGGGGGCTGGAGCACCA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TGGGGTGACGGCAGGGCAGGGGGCGCCTGGCCGGGGAGAAGCGCGGGGGCTGGAGCACCA 120

Qy        121 CCAACTGGAGGGTCCGGAGTAGCGAGCGCCCCGAAGGAGGCCATCGGGGAGCCGGGAGGG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CCAACTGGAGGGTCCGGAGTAGCGAGCGCCCCGAAGGAGGCCATCGGGGAGCCGGGAGGG 180

Qy        181 GGGACTGCGAGAGGACCCCGGCGTCCGGGCTCCCGGTGCCAGCGCTATGAGGCCACTCCT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGGACTGCGAGAGGACCCCGGCGTCCGGGCTCCCGGTGCCAGCGCTATGAGGCCACTCCT 240

Qy        241 CGTCCTGCTGCTCCTGGGCCTGGCGGCCGGCTCGCCCCCACTGGACGACAACAAGATCCC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CGTCCTGCTGCTCCTGGGCCTGGCGGCCGGCTCGCCCCCACTGGACGACAACAAGATCCC 300

Qy        301 CAGCCTCTGCCCGGGGCACCCCGGCCTTCCAGGCACGCCGGGCCACCATGGCAGCCAGGG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CAGCCTCTGCCCGGGGCACCCCGGCCTTCCAGGCACGCCGGGCCACCATGGCAGCCAGGG 360

Qy        361 CTTGCCGGGCCGCGATGGCCGCGACGGCCGCGACGGCGCGCCCGGGGCTCCGGGAGAGAA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CTTGCCGGGCCGCGATGGCCGCGACGGCCGCGACGGCGCGCCCGGGGCTCCGGGAGAGAA 420

Qy        421 AGGCGAGGGCGGGAGGCCGGGACTGCCGGGACCTCGAGGGGACCCCGGGCCGCGAGGAGA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AGGCGAGGGCGGGAGGCCGGGACTGCCGGGACCTCGAGGGGACCCCGGGCCGCGAGGAGA 480

Qy        481 GGCGGGACCCGCGGGGGCCCACCGGGCCTGCCGGGGAGTGCTCGGTGCCTCCGCGATCCGC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GGCGGGACCCGCGGGGGCCCACCGGGCCTGCCGGGGAGTGCTCGGTGCCTCCGCGATCCGC 540

Qy        541 CTTCAGCGCCAAGCGCTCCGAGAGCCGGGTGCCTCCGCCGTCTGACGCACCCTTGCCCTT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CTTCAGCGCCAAGCGCTCCGAGAGCCGGGTGCCTCCGCCGTCTGACGCACCCTTGCCCTT 600

Qy        601 CGACCGCGTGCTGGTGAACGAGCAGGGACATTACGACGCCGTCACCGGCAAGTTCACCTG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CGACCGCGTGCTGGTGAACGAGCAGGGACATTACGACGCCGTCACCGGCAAGTTCACCTG 660

Qy        661 CCAGGTGCCTGGGGTCTACTACTTCGCCGTCCATGCCACCGTCTACCGGGCCAGCCTGCA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CCAGGTGCCTGGGGTCTACTACTTCGCCGTCCATGCCACCGTCTACCGGGCCAGCCTGCA 720

Qy        721 GTTTGATCTGGTGAAGAATGGCGAATCCATTGCCTCTTTCTTCCAGTTTTTCGGGGGGTG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GTTTGATCTGGTGAAGAATGGCGAATCCATTGCCTCTTTCTTCCAGTTTTTCGGGGGGTG 780

Qy        781 GCCCAAGCCAGCCTCGCTCTCGGGGGGGGCCATGGTGAGGCTGGAGCCTGAGGACCAAGT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GCCCAAGCCAGCCTCGCTCTCGGGGGGGGCCATGGTGAGGCTGGAGCCTGAGGACCAAGT 840

Qy        841 GTGGGTGCAGGTGGGTGTGGGTGACTACATTGGCATCTATGCCAGCATCAAGACAGACAG 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GTGGGTGCAGGTGGGTGTGGGTGACTACATTGGCATCTATGCCAGCATCAAGACAGACAG 900

Qy        901 CACCTTCTCCGGATTTCTGGTGTACTCCGACTGGCACAGCTCCCCAGTCTTTGCTTAGTG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 CACCTTCTCCGGATTTCTGGTGTACTCCGACTGGCACAGCTCCCCAGTCTTTGCTTAGTG 960

Qy        961 CCCACTGCAAAGTGAGCTCATGCTCTCACTCCTAGAAGGAGGGTGTGAGGCTGACAACCA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 CCCACTGCAAAGTGAGCTCATGCTCTCACTCCTAGAAGGAGGGTGTGAGGCTGACAACCA 1020

Qy       1021 GGTCATCCAGGAGGGCTGGCCCCCCTGGAATATTGTGAATGACTAGGGAGGTGGGGTAGA 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GGTCATCCAGGAGGGCTGGCCCCCCTGGAATATTGTGAATGACTAGGGAGGTGGGGTAGA 1080

Qy       1081 GCACTCTCCGTCCTGCTGCTGGCAAGGAATGGGAACAGTGGCTGTCTGCGATCAGGTCTG 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 GCACTCTCCGTCCTGCTGCTGGCAAGGAATGGGAACAGTGGCTGTCTGCGATCAGGTCTG 1140

Qy       1141 GCAGCATGGGGCAGTGGCTGGATTTCTGCCCAAGACCAGAGGAGTGTGCTGTGCTGGCAA 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 GCAGCATGGGGCAGTGGCTGGATTTCTGCCCAAGACCAGAGGAGTGTGCTGTGCTGGCAA 1200

Qy       1201 GTGTAAGTCCCCCAGTTGCTCTGGTCCAGGAGCCCACGGTGGGGTGCTCTCTTCCTGGTC 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 GTGTAAGTCCCCCAGTTGCTCTGGTCCAGGAGCCCACGGTGGGGTGCTCTCTTCCTGGTC 1260

Qy       1261 CTCTGCTTCTCTGGATCCTCCCCACCCCCTCCTGCTCCTGGGGCCGGCCCTTTTCTCAGA 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 CTCTGCTTCTCTGGATCCTCCCCACCCCCTCCTGCTCCTGGGGCCGGCCCTTTTCTCAGA 1320

Qy       1321 GATCACTCAATAAACCTAAGAACCCTCATAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1377
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 GATCACTCAATAAACCTAAGAACCCTCATAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1377
```

RESULT 3
AAD01241
ID AAD01241 standard; cDNA; 1377 BP.
XX
AC AAD01241;
XX

```
DT   08-NOV-2000  (first entry)
XX
DE   Human PRO344 protein encoding cDNA clone, DNA40592-1242.
XX
KW   PRO344; DNA40592-1242; human; ATCC No: 209492; antiproliferative;
KW   neoplastic cell growth inhibitor; cytostatic; treatment; cancer; tumour;
KW   breast; prostate; colon; lung; renal; ovarian; central nervous system;
KW   CNS; leukemia; melanoma; Expressed Sequence Tag; EST; secreted protein;
KW   extracellular domain; ECD; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             227..958
FT                   /*tag= a
FT                   /product= "Human PRO344 protein"
FT                   /note= "Derived from clone DNA40592-1242"
FT   sig_peptide     227..271
FT                   /*tag= b
FT   mat_peptide     272..955
FT                   /*tag= c
FT                   /product= "Mature human PRO344 protein"
XX
PN   WO200032778-A2.
XX
PD   08-JUN-2000.
XX
PF   30-NOV-1999;   99WO-US28409.
XX
PR   01-DEC-1998;   98WO-US25108.
PR   16-DEC-1998;   98US-0112850.
PR   22-DEC-1998;   98US-0113296.
PR   20-JUL-1999;   99US-0144758.
PR   26-JUL-1999;   99US-0145698.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Chen J,  Goddard A,  Gurney AL,  Hillan K,  Napier M,  Wood WI;
XX
DR   WPI; 2000-412325/35.
DR   P-PSDB; AAY71468.
XX
PT   New composition useful for inhibiting neoplastic cell growth and for
PT   treating cancers, comprises PRO655, PRO344 or PRO364 polypeptide or
PT   their antagonists -
XX
PS   Claim 20; Fig 5; 108pp; English.
XX
CC   The present sequence is the cDNA clone, designated as DNA40592-1242,
CC   encoding the human PRO344 polypeptide. It is isolated from human foetal
CC   lung tissue cDNA library, identified using probes based on a consensus
CC   sequence DNA34398, derived from secreted protein extracellular domain
CC   (ECD) Expressed Sequence Tag (EST). This clone is assigned ATCC deposit
CC   No: 209492. PRO344 functions as a neoplastic cell growth inhibitor and
CC   is used for treating tumours, using an effective amount of PRO655, PRO364
CC   and PRO344. This composition is especially useful for treatment of human
CC   cancers such as breast, prostate, colon, lung, renal, ovarian and CNS,
CC   leukemia and melanoma.
XX
SQ   Sequence 1377 BP; 251 A; 422 C; 472 G; 232 T; 0 other;
```

Query Match 100.0%; Score 1377; DB 21; Length 1377;
Best Local Similarity 100.0%; Pred. No. 1.7e-244;
Matches 1377; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 GACTAGTTCTCTTGGAGTCTGGGAGGAGGAAAGCGGAGCCGGCAGGGAGCGAACCAGGAC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GACTAGTTCTCTTGGAGTCTGGGAGGAGGAAAGCGGAGCCGGCAGGGAGCGAACCAGGAC 60

Qy         61 TGGGGTGACGGCAGGGCAGGGGGCGCCTGGCCGGGGAGAAGCGCGGGGGCTGGAGCACCA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TGGGGTGACGGCAGGGCAGGGGGCGCCTGGCCGGGGAGAAGCGCGGGGGCTGGAGCACCA 120

Qy        121 CCAACTGGAGGGTCCGGAGTAGCGAGCGCCCCGAAGGAGGCCATCGGGGAGCCGGGAGGG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CCAACTGGAGGGTCCGGAGTAGCGAGCGCCCCGAAGGAGGCCATCGGGGAGCCGGGAGGG 180

Qy        181 GGGACTGCGAGAGGACCCCGGCGTCCGGGCTCCCGGTGCCAGCGCTATGAGGCCACTCCT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGGACTGCGAGAGGACCCCGGCGTCCGGGCTCCCGGTGCCAGCGCTATGAGGCCACTCCT 240

Qy        241 CGTCCTGCTGCTCCTGGGCCTGGCGGCCGGCTCGCCCCCACTGGACGACAACAAGATCCC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CGTCCTGCTGCTCCTGGGCCTGGCGGCCGGCTCGCCCCCACTGGACGACAACAAGATCCC 300

Qy        301 CAGCCTCTGCCCGGGGCACCCCGGCCTTCCAGGCACGCCGGGCCACCATGGCAGCCAGGG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CAGCCTCTGCCCGGGGCACCCCGGCCTTCCAGGCACGCCGGGCCACCATGGCAGCCAGGG 360

Qy        361 CTTGCCGGGCCGCGATGGCCGCGACGGCCGCGACGGCGCGCCCGGGGCTCCGGGAGAGAA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CTTGCCGGGCCGCGATGGCCGCGACGGCCGCGACGGCGCGCCCGGGGCTCCGGGAGAGAA 420

Qy        421 AGGCGAGGGCGGGAGGCCGGGACTGCCGGGACCTCGAGGGGACCCCGGGCCGCGAGGAGA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AGGCGAGGGCGGGAGGCCGGGACTGCCGGGACCTCGAGGGGACCCCGGGCCGCGAGGAGA 480

Qy        481 GGCGGGACCCGCGGGGCCCACCGGGCCTGCCGGGGAGTGCTCGGTGCCTCCGCGATCCGC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GGCGGGACCCGCGGGGCCCACCGGGCCTGCCGGGGAGTGCTCGGTGCCTCCGCGATCCGC 540

Qy        541 CTTCAGCGCCAAGCGCTCCGAGAGCCGGGTGCCTCCGCCGTCTGACGCACCCTTGCCCTT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CTTCAGCGCCAAGCGCTCCGAGAGCCGGGTGCCTCCGCCGTCTGACGCACCCTTGCCCTT 600

Qy        601 CGACCGCGTGCTGGTGAACGAGCAGGGACATTACGACGCCGTCACCGGCAAGTTCACCTG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CGACCGCGTGCTGGTGAACGAGCAGGGACATTACGACGCCGTCACCGGCAAGTTCACCTG 660

Qy        661 CCAGGTGCCTGGGGTCTACTACTTCGCCGTCCATGCCACCGTCTACCGGGCCAGCCTGCA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CCAGGTGCCTGGGGTCTACTACTTCGCCGTCCATGCCACCGTCTACCGGGCCAGCCTGCA 720

Qy        721 GTTTGATCTGGTGAAGAATGGCGAATCCATTGCCTCTTTCTTCCAGTTTTTCGGGGGGTG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GTTTGATCTGGTGAAGAATGGCGAATCCATTGCCTCTTTCTTCCAGTTTTTCGGGGGGTG 780

Qy        781 GCCCAAGCCAGCCTCGCTCTCGGGGGGGGCCATGGTGAGGCTGGAGCCTGAGGACCAAGT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GCCCAAGCCAGCCTCGCTCTCGGGGGGGGCCATGGTGAGGCTGGAGCCTGAGGACCAAGT 840

Qy        841 GTGGGTGCAGGTGGGTGTGGGTGACTACATTGGCATCTATGCCAGCATCAAGACAGACAG 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GTGGGTGCAGGTGGGTGTGGGTGACTACATTGGCATCTATGCCAGCATCAAGACAGACAG 900

Qy        901 CACCTTCTCCGGATTTCTGGTGTACTCCGACTGGCACAGCTCCCCAGTCTTTGCTTAGTG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 CACCTTCTCCGGATTTCTGGTGTACTCCGACTGGCACAGCTCCCCAGTCTTTGCTTAGTG 960

Qy        961 CCCACTGCAAAGTGAGCTCATGCTCTCACTCCTAGAAGGAGGGTGTGAGGCTGACAACCA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 CCCACTGCAAAGTGAGCTCATGCTCTCACTCCTAGAAGGAGGGTGTGAGGCTGACAACCA 1020

Qy       1021 GGTCATCCAGGAGGGCTGGCCCCCCTGGAATATTGTGAATGACTAGGGAGGTGGGGTAGA 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GGTCATCCAGGAGGGCTGGCCCCCCTGGAATATTGTGAATGACTAGGGAGGTGGGGTAGA 1080

Qy       1081 GCACTCTCCGTCCTGCTGCTGGCAAGGAATGGGAACAGTGGCTGTCTGCGATCAGGTCTG 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 GCACTCTCCGTCCTGCTGCTGGCAAGGAATGGGAACAGTGGCTGTCTGCGATCAGGTCTG 1140

Qy       1141 GCAGCATGGGGCAGTGGCTGGATTTCTGCCCAAGACCAGAGGAGTGTGCTGTGCTGGCAA 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 GCAGCATGGGGCAGTGGCTGGATTTCTGCCCAAGACCAGAGGAGTGTGCTGTGCTGGCAA 1200

Qy       1201 GTGTAAGTCCCCCAGTTGCTCTGGTCCAGGAGCCCACGGTGGGGTGCTCTCTTCCTGGTC 1260
```

```
Db       1201 GTGTAAGTCCCCCAGTTGCTCTGGTCCAGGAGCCCACGGTGGGGTGCTCTCTTCCTGGTC 1260

Qy       1261 CTCTGCTTCTCTGGATCCTCCCCACCCCCTCCTGCTCCTGGGGCCGGCCCTTTTCTCAGA 1320

Db       1261 CTCTGCTTCTCTGGATCCTCCCCACCCCCTCCTGCTCCTGGGGCCGGCCCTTTTCTCAGA 1320

Qy       1321 GATCACTCAATAAACCTAAGAACCCTCATAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1377

Db       1321 GATCACTCAATAAACCTAAGAACCCTCATAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1377
```

RESULT 4
AAA46907
ID AAA46907 standard; cDNA; 1377 BP.
XX
AC AAA46907;
XX
DT 03-OCT-2000 (first entry)
XX
DE cDNA encoding novel polypeptide PRO344.
XX
KW PRO201; PRO292; PRO327; PRO1265; PRO344; PRO343; PRO347; PRO357;
KW PRO715; PRO1017; PRO1112; PRO509; PRO853; PRO882; tumour cell;
KW tumourigenesis; cancer; neoplastic cell growth; cell proliferation; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 227..958
FT /*tag= a
XX
PN WO200037640-A2.
XX
PD 29-JUN-2000.
XX
PF 16-DEC-1999; 99WO-US30095.
XX
PR 22-DEC-1998; 98US-0113296.
PR 08-MAR-1999; 99WO-US05028.
PR 02-JUN-1999; 99WO-US12252.
PR 01-SEP-1999; 99WO-US20111.
PR 15-SEP-1999; 99WO-US21090.
PR 30-NOV-1999; 99WO-US28313.
PR 30-NOV-1999; 99WO-US28409.
PR 01-DEC-1999; 99WO-US28301.
PR 02-DEC-1999; 99WO-US28565.
XX
PA (GETH ) GENENTECH INC.
XX
PI Botstein D, Goddard A, Gurney AL, Hillan K, Lawrence DA, Roy MA;
PI Wood WI;
XX
DR WPI; 2000-452188/39.
DR P-PSDB; AAY93688.
XX
PT New anti-polypeptide antibody useful in the treatment and diagnosis of
PT neoplastic cell growth and proliferation -
XX
PS Claim 50; Fig 9; 220pp; English.
XX
CC The present sequence encodes a novel human polypeptide. The
CC specification describes novel polypeptides designated PRO201, PRO292,
CC PRO327, PRO1265, PRO344, PRO343, PRO347, PRO357, PRO715, PRO1017,
CC PRO1112, PRO509, PRO853 and PRO882. These genes are amplified in
CC the genome of tumour cells. The polypeptides are believed to contribute
CC to tumourigenesis. The polypeptides are useful target for the
CC identification of certain cancers, and may act as predictors of the
CC prognosis of tumour treatment. Antibodies against these polypeptides
CC are useful in the treatment and diagnosis of neoplastic cell growth
CC and proliferation in mammals.
XX
SQ Sequence 1377 BP; 251 A; 422 C; 472 G; 232 T; 0 other;

Query Match 100.0%; Score 1377; DB 21; Length 1377;
Best Local Similarity 100.0%; Pred. No. 1.7e-244;
Matches 1377; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 GACTAGTTCTCTTGGAGTCTGGGAGGAGGAAAGCGGAGCCGGCAGGGAGCGAACCAGGAC 60

Db          1 GACTAGTTCTCTTGGAGTCTGGGAGGAGGAAAGCGGAGCCGGCAGGGAGCGAACCAGGAC 60

Qy         61 TGGGGTGACGGCAGGGCAGGGGGCGCCTGGCCGGGGAGAAGCGCGGGGGCTGGAGCACCA 120

Db         61 TGGGGTGACGGCAGGGCAGGGGGCGCCTGGCCGGGGAGAAGCGCGGGGGCTGGAGCACCA 120

Qy        121 CCAACTGGAGGGTCCGGAGTAGCGAGCGCCCCGAAGGAGGCCATCGGGGAGCCGGGAGGG 180

Db        121 CCAACTGGAGGGTCCGGAGTAGCGAGCGCCCCGAAGGAGGCCATCGGGGAGCCGGGAGGG 180

Qy        181 GGGACTGCGAGAGGACCCCGGCGTCCGGGCTCCCGGTGCCAGCGCTATGAGGCCACTCCT 240

Db        181 GGGACTGCGAGAGGACCCCGGCGTCCGGGCTCCCGGTGCCAGCGCTATGAGGCCACTCCT 240

Qy        241 CGTCCTGCTGCTCCTGGGCCTGGCGGCCGGCTCGCCCCCACTGGACGACAACAAGATCCC 300

Db        241 CGTCCTGCTGCTCCTGGGCCTGGCGGCCGGCTCGCCCCCACTGGACGACAACAAGATCCC 300

Qy        301 CAGCCTCTGCCCGGGGCACCCCGGCCTTCCAGGCACGCCGGGCCACCATGGCAGCCAGGG 360

Db        301 CAGCCTCTGCCCGGGGCACCCCGGCCTTCCAGGCACGCCGGGCCACCATGGCAGCCAGGG 360

Qy        361 CTTGCCGGGCCGCGATGGCCGCGACGGCCGCGACGGCGCGCCCGGGGCTCCGGGAGAGAA 420

Db        361 CTTGCCGGGCCGCGATGGCCGCGACGGCCGCGACGGCGCGCCCGGGGCTCCGGGAGAGAA 420

Qy        421 AGGCGAGGGCGGGAGGCCGGGACTGCCGGGACCTCGAGGGGACCCCGGGCCGCGAGGAGA 480

Db        421 AGGCGAGGGCGGGAGGCCGGGACTGCCGGGACCTCGAGGGGACCCCGGGCCGCGAGGAGA 480

Qy        481 GGCGGGACCCGCGGGGCCCACCGGGCCTGCCGGGGAGTGCTCGGTGCCTCCGCGATCCGC 540

Db        481 GGCGGGACCCGCGGGGCCCACCGGGCCTGCCGGGGAGTGCTCGGTGCCTCCGCGATCCGC 540

Qy        541 CTTCAGCGCCAAGCGCTCCGAGAGCCGGGTGCCTCCGCCGTCTGACGCACCCTTGCCCTT 600

Db        541 CTTCAGCGCCAAGCGCTCCGAGAGCCGGGTGCCTCCGCCGTCTGACGCACCCTTGCCCTT 600

Qy        601 CGACCGCGTGCTGGTGAACGAGCAGGGACATTACGACGCCGTCACCGGCAAGTTCACCTG 660

Db        601 CGACCGCGTGCTGGTGAACGAGCAGGGACATTACGACGCCGTCACCGGCAAGTTCACCTG 660

Qy        661 CCAGGTGCCTGGGGTCTACTACTTCGCCGTCCATGCCACCGTCTACCGGGCCAGCCTGCA 720

Db        661 CCAGGTGCCTGGGGTCTACTACTTCGCCGTCCATGCCACCGTCTACCGGGCCAGCCTGCA 720

Qy        721 GTTTGATCTGGTGAAGAATGGCGAATCCATTGCCTCTTTCTTCCAGTTTTTCGGGGGGTG 780

Db        721 GTTTGATCTGGTGAAGAATGGCGAATCCATTGCCTCTTTCTTCCAGTTTTTCGGGGGGTG 780

Qy        781 GCCCAAGCCAGCCTCGCTCTCGGGGGGGGCCATGGTGAGGCTGGAGCCTGAGGACCAAGT 840

Db        781 GCCCAAGCCAGCCTCGCTCTCGGGGGGGGCCATGGTGAGGCTGGAGCCTGAGGACCAAGT 840

Qy        841 GTGGGTGCAGGTGGGTGTGGGTGACTACATTGGCATCTATGCCAGCATCAAGACAGACAG 900

Db        841 GTGGGTGCAGGTGGGTGTGGGTGACTACATTGGCATCTATGCCAGCATCAAGACAGACAG 900

Qy        901 CACCTTCTCCGGATTTCTGGTGTACTCCGACTGGCACAGCTCCCCAGTCTTTGCTTAGTG 960

Db        901 CACCTTCTCCGGATTTCTGGTGTACTCCGACTGGCACAGCTCCCCAGTCTTTGCTTAGTG 960

Qy        961 CCCACTGCAAAGTGAGCTCATGCTCTCACTCCTAGAAGGAGGGTGTGAGGCTGACAACCA 1020

Db        961 CCCACTGCAAAGTGAGCTCATGCTCTCACTCCTAGAAGGAGGGTGTGAGGCTGACAACCA 1020
```

```
Qy       1021 GGTCATCCAGGAGGGCTGGCCCCCCTGGAATATTGTGAATGACTAGGGAGGTGGGGTAGA 1080
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GGTCATCCAGGAGGGCTGGCCCCCCTGGAATATTGTGAATGACTAGGGAGGTGGGGTAGA 1080

Qy       1081 GCACTCTCCGTCCTGCTGCTGGCAAGGAATGGGAACAGTGGCTGTCTGCGATCAGGTCTG 1140
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 GCACTCTCCGTCCTGCTGCTGGCAAGGAATGGGAACAGTGGCTGTCTGCGATCAGGTCTG 1140

Qy       1141 GCAGCATGGGGCAGTGGCTGGATTTCTGCCCAAGACCAGAGGAGTGTGCTGTGCTGGCAA 1200
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 GCAGCATGGGGCAGTGGCTGGATTTCTGCCCAAGACCAGAGGAGTGTGCTGTGCTGGCAA 1200

Qy       1201 GTGTAAGTCCCCCAGTTGCTCTGGTCCAGGAGCCCACGGTGGGGTGCTCTCTTCCTGGTC 1260
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 GTGTAAGTCCCCCAGTTGCTCTGGTCCAGGAGCCCACGGTGGGGTGCTCTCTTCCTGGTC 1260

Qy       1261 CTCTGCTTCTCTGGATCCTCCCCACCCCCTCCTGCTCCTGGGGCCGGCCCTTTTCTCAGA 1320
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 CTCTGCTTCTCTGGATCCTCCCCACCCCCTCCTGCTCCTGGGGCCGGCCCTTTTCTCAGA 1320

Qy       1321 GATCACTCAATAAACCTAAGAACCCTCATAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1377
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 GATCACTCAATAAACCTAAGAACCCTCATAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1377
```

```
RESULT 5
AAA49560
ID   AAA49560 standard; cDNA; 1377 BP.
XX
AC   AAA49560;
XX
DT   25-SEP-2000  (first entry)
XX
DE   Human PRO344 cDNA.
XX
KW   PRO; membrane bound protein; secreted protein; PRO357; PRO327;
KW   PRO243; PRO715; PRO241; PRO323; PRO299; PRO233; PRO344; PRO347;
KW   PRO355; PRO353; PRO361; PRO365; transmembrane polypeptide;
KW   antibody; screening; detection; inhibition; probe; primer; human;
KW   ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             227..958
FT                   /*tag= a
FT                   /product= PRO344 polypeptide
XX
PN   WO200032776-A2.
XX
PD   08-JUN-2000.
XX
PF   01-DEC-1999;   99WO-US28301.
XX
PR   01-DEC-1998;   98WO-US25108.
PR   16-DEC-1998;   98US-0112850.
PR   22-DEC-1998;   98US-0113296.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Baker KP,  Botstein D,  Eaton DL,  Ferrara N,  Filvaroff E;
PI   Gerritsen ME,  Goddard A,  Godowski PJ,  Grimaldi CJ,  Gurney AL;
PI   Hillan KJ,  Kljavin IJ,  Napier MA,  Roy MA,  Tumas D,  Wood WI;
XX
DR   WPI; 2000-412324/35.
DR   P-PSDB; AAB01318.
XX
PT   New human nucleic acids encoding secreted and transmembrane
PT   polypeptides, designated as PRO polypeptides, useful as pharmaceutical
PT   and diagnostic agents
XX
PS   Claim 2; Fig 17; 187pp; English.
XX
CC   New human nucleic acids encoding secreted and transmembrane
CC   polypeptides which are designated as PRO polypeptides are described
CC   The membrane-bound proteins have various industrial applications,
CC   including as pharmaceutical and diagnostic agents. The membrane-bound
CC   proteins can also be employed for screening of potential peptide or
CC   small molecule inhibitors of the relevant receptor/ligand interaction.
CC   Anti-PRO antibodies are useful for the affinity purification of PRO
CC   from recombinant cell culture or natural sources.
XX
SQ   Sequence 1377 BP; 251 A; 422 C; 472 G; 232 T; 0 other;
```

```
  Query Match           100.0%;  Score 1377;  DB 21;  Length 1377;
  Best Local Similarity  100.0%;  Pred. No. 1.7e-244;
  Matches 1377;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GACTAGTTCTCTTGGAGTCTGGGAGGAGGAAAGCGGAGCCGGCAGGGAGCGAACCAGGAC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GACTAGTTCTCTTGGAGTCTGGGAGGAGGAAAGCGGAGCCGGCAGGGAGCGAACCAGGAC 60

Qy         61 TGGGGTGACGGCAGGGCAGGGGGCGCCTGGCCGGGGAGAAGCGCGGGGGCTGGAGCACCA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TGGGGTGACGGCAGGGCAGGGGGCGCCTGGCCGGGGAGAAGCGCGGGGGCTGGAGCACCA 120

Qy        121 CCAACTGGAGGGTCCGGAGTAGCGAGCGCCCCGAAGGAGGCCATCGGGGAGCCGGGAGGG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CCAACTGGAGGGTCCGGAGTAGCGAGCGCCCCGAAGGAGGCCATCGGGGAGCCGGGAGGG 180

Qy        181 GGGACTGCGAGAGGACCCCGGCGTCCGGGCTCCCGGTGCCAGCGCTATGAGGCCACTCCT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGGACTGCGAGAGGACCCCGGCGTCCGGGCTCCCGGTGCCAGCGCTATGAGGCCACTCCT 240

Qy        241 CGTCCTGCTGCTCCTGGGCCTGGCGGCCGGCTCGCCCCCACTGGACGACAACAAGATCCC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CGTCCTGCTGCTCCTGGGCCTGGCGGCCGGCTCGCCCCCACTGGACGACAACAAGATCCC 300

Qy        301 CAGCCTCTGCCCGGGGCACCCCGGCCTTCCAGGCACGCCGGGCCACCATGGCAGCCAGGG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CAGCCTCTGCCCGGGGCACCCCGGCCTTCCAGGCACGCCGGGCCACCATGGCAGCCAGGG 360

Qy        361 CTTGCCGGGCCGCGATGGCCGCGACGGCCGCGACGGCGCGCCCGGGGCTCCGGGAGAGAA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CTTGCCGGGCCGCGATGGCCGCGACGGCCGCGACGGCGCGCCCGGGGCTCCGGGAGAGAA 420

Qy        421 AGGCGAGGGCGGGAGGCCGGGACTGCCGGGACCTCGAGGGGACCCCGGGCCGCGAGGAGA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AGGCGAGGGCGGGAGGCCGGGACTGCCGGGACCTCGAGGGGACCCCGGGCCGCGAGGAGA 480

Qy        481 GGCGGGACCCGCGGGGGCCCACCGGGCCTGCCGGGGAGTGCTCGGTGCCTCCGCGATCCGC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GGCGGGACCCGCGGGGGCCCACCGGGCCTGCCGGGGAGTGCTCGGTGCCTCCGCGATCCGC 540

Qy        541 CTTCAGCGCCAAGCGCTCCGAGAGCCGGGTGCCTCCGCCGTCTGACGCACCCTTGCCCTT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CTTCAGCGCCAAGCGCTCCGAGAGCCGGGTGCCTCCGCCGTCTGACGCACCCTTGCCCTT 600

Qy        601 CGACCGCGTGCTGGTGAACGAGCAGGGACATTACGACGCCGTCACCGGCAAGTTCACCTG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CGACCGCGTGCTGGTGAACGAGCAGGGACATTACGACGCCGTCACCGGCAAGTTCACCTG 660

Qy        661 CCAGGTGCCTGGGGTCTACTACTTCGCCGTCCATGCCACCGTCTACCGGGCCAGCCTGCA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CCAGGTGCCTGGGGTCTACTACTTCGCCGTCCATGCCACCGTCTACCGGGCCAGCCTGCA 720

Qy        721 GTTTGATCTGGTGAAGAATGGCGAATCCATTGCCTCTTTCTTCCAGTTTTTCGGGGGGTG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GTTTGATCTGGTGAAGAATGGCGAATCCATTGCCTCTTTCTTCCAGTTTTTCGGGGGGTG 780

Qy        781 GCCCAAGCCAGCCTCGCTCTCGGGGGGGGCCATGGTGAGGCTGGAGCCTGAGGACCAAGT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GCCCAAGCCAGCCTCGCTCTCGGGGGGGGCCATGGTGAGGCTGGAGCCTGAGGACCAAGT 840

Qy        841 GTGGGTGCAGGTGGGTGTGGGTGACTACATTGGCATCTATGCCAGCATCAAGACAGACAG 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        841 GTGGGTGCAGGTGGGTGTGGGTGACTACATTGGCATCTATGCCAGCATCAAGACAGACAG 900
Qy        901 CACCTTCTCCGGATTTCTGGTGTACTCCGACTGGCACAGCTCCCCAGTCTTTGCTTAGTG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 CACCTTCTCCGGATTTCTGGTGTACTCCGACTGGCACAGCTCCCCAGTCTTTGCTTAGTG 960
Qy        961 CCCACTGCAAAGTGAGCTCATGCTCTCACTCCTAGAAGGAGGGTGTGAGGCTGACAACCA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 CCCACTGCAAAGTGAGCTCATGCTCTCACTCCTAGAAGGAGGGTGTGAGGCTGACAACCA 1020
Qy       1021 GGTCATCCAGGAGGGCTGGCCCCCCTGGAATATTGTGAATGACTAGGGAGGTGGGGTAGA 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GGTCATCCAGGAGGGCTGGCCCCCCTGGAATATTGTGAATGACTAGGGAGGTGGGGTAGA 1080
Qy       1081 GCACTCTCCGTCCTGCTGCTGGCAAGGAATGGGAACAGTGGCTGTCTGCGATCAGGTCTG 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 GCACTCTCCGTCCTGCTGCTGGCAAGGAATGGGAACAGTGGCTGTCTGCGATCAGGTCTG 1140
Qy       1141 GCAGCATGGGGCAGTGGCTGGATTTCTGCCCAAGACCAGAGGAGTGTGCTGTGCTGGCAA 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 GCAGCATGGGGCAGTGGCTGGATTTCTGCCCAAGACCAGAGGAGTGTGCTGTGCTGGCAA 1200
Qy       1201 GTGTAAGTCCCCCAGTTGCTCTGGTCCAGGAGCCCACGGTGGGGTGCTCTCTTCCTGGTC 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 GTGTAAGTCCCCCAGTTGCTCTGGTCCAGGAGCCCACGGTGGGGTGCTCTCTTCCTGGTC 1260
Qy       1261 CTCTGCTTCTCTGGATCCTCCCCACCCCCTCCTGCTCCTGGGGCCGGCCCTTTTCTCAGA 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 CTCTGCTTCTCTGGATCCTCCCCACCCCCTCCTGCTCCTGGGGCCGGCCCTTTTCTCAGA 1320
Qy       1321 GATCACTCAATAAACCTAAGAACCCTCATAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1377
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 GATCACTCAATAAACCTAAGAACCCTCATAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1377

RESULT 6
AAS21424
ID   AAS21424 standard; cDNA; 1377 BP.
XX
AC   AAS21424;
XX
DT   24-OCT-2001  (first entry)
XX
DE   Human cDNA sequence encoding for PRO344 polypeptide.
XX
KW   Human secretory and transmembrane; PRO; mammalian; cancer; lung;
KW   breast; prostate; cervical; tumour necrosis factor-alpha; TNF-alpha;
KW   cartilage; ear; proliferation; glucose; free fatty acid; skeletal muscle;
KW   adipocyte; A-peptide; factor VIIA; gene therapy; ss.
XX
OS   Homo sapiens.
XX
PN   WO200140466-A2.
XX
PD   07-JUN-2001.
XX
PF   01-DEC-2000; 2000WO-US32678.
XX
PR   01-DEC-1999;   99WO-US28301.
PR   01-DEC-1999;   99WO-US28634.
PR   02-DEC-1999;   99WO-US28551.
PR   02-DEC-1999;   99WO-US28564.
PR   02-DEC-1999;   99WO-US28565.
PR   09-DEC-1999;   99US-0170262.
PR   16-DEC-1999;   99WO-US30095.
PR   20-DEC-1999;   99WO-US30911.
PR   20-DEC-1999;   99WO-US30999.
PR   30-DEC-1999;   99WO-US31243.
PR   06-JAN-2000; 2000WO-US00277.
PR   06-JAN-2000; 2000WO-US00376.
PR   11-FEB-2000; 2000WO-US03565.
PR   18-FEB-2000; 2000WO-US04341.
PR   18-FEB-2000; 2000WO-US04342.
PR   22-FEB-2000; 2000WO-US04414.
PR   24-FEB-2000; 2000WO-US04914.
PR   24-FEB-2000; 2000WO-US05004.
PR   01-MAR-2000; 2000WO-US05601.
PR   20-MAR-2000; 2000WO-US07377.
PR   21-MAR-2000; 2000WO-US07532.
PR   30-MAR-2000; 2000WO-US08439.
PR   17-MAY-2000; 2000WO-US13705.
PR   22-MAY-2000; 2000WO-US14042.
PR   30-MAY-2000; 2000WO-US14941.
PR   02-JUN-2000; 2000WO-US15264.
PR   10-NOV-2000; 2000WO-US30873.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Baker KP,  Beresini M,  Deforge L,  Desnoyers L,  Filvaroff E,  Gao W;
PI   Gerritsen ME,  Goddard A,  Godowski PJ,  Gurney AL,  Sherwood S;
PI   Smith V,  Stewart TA,  Tumas D,  Watanabe CK,  Wood WI,  Zhang Z;
XX
DR   WPI; 2001-408281/43.
DR   P-PSDB; AAU12352.
XX
PT   Isolated, secretory and transmembrane PRO polypeptide used to detect
PT   other PRO polypeptides, link bioactive molecules to cells expressing
PT   PRO polypeptides, and detect the presence of mammalian tumours e.g.
PT   lung, breast, prostate, cervical -
XX
PS   Claim 3; Fig 361; 813pp; English.
XX
CC   AAS21244-AAS21518 encode for novel human secretory and transmembrane
CC   PRO polypeptides. The PRO polypeptides are useful to detect other
CC   PRO polypeptides, to link bioactive molecules to cells expressing
CC   PRO polypeptides, to modulate biological activities of cells expressing
CC   PRO polypeptides, and to detect the presence of mammalian lung, colon,
CC   breast, prostate, rectal, cervical or liver tumours by comparing PRO
CC   polypeptide expression in a cell sample to that in a control sample.
CC   Some of the 275 sequences are also useful to stimulate the release of
CC   tumour necrosis factor-alpha (TNF-alpha) from human blood, the
CC   proliferation or differentiation of chondrocytes, the proliferation or
CC   gene expression in pericyte cells, the release of proteoglycans from
CC   cartilage, the proliferation of inner ear utricular supporting cells or
CC   of T-lymphocytes, the release of a cytokine from peripheral blood
CC   monocytes (PBMCs), or the proliferation of endothelial cells. Some of
CC   the PRO polypeptides may modulate glucose or free fatty acid uptake by
CC   skeletal muscle cells or by adipocytes; or inhibit binding of A-peptide
CC   to factor VIIA. The PRO polypeptides can be used in assays to identify
CC   molecules involved in binding interactions. The polynucleotides encoding
CC   PRO polypeptides can be used to generate probes, antisense RNA/DNA,
CC   transgenic or knock out animals and can be used in gene therapy.
XX
SQ   Sequence 1377 BP; 251 A; 422 C; 472 G; 232 T; 0 other;

  Query Match           100.0%;  Score 1377;  DB 22;  Length 1377;
  Best Local Similarity  100.0%;  Pred. No. 1.7e-244;
  Matches 1377;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GACTAGTTCTCTTGGAGTCTGGGAGGAGGAAAGCGGAGCCGGCAGGGAGCGAACCAGGAC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GACTAGTTCTCTTGGAGTCTGGGAGGAGGAAAGCGGAGCCGGCAGGGAGCGAACCAGGAC 60
Qy         61 TGGGGTGACGGCAGGGCAGGGGGCGCCTGGCCGGGGAGAAGCGCGGGGGCTGGAGCACCA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TGGGGTGACGGCAGGGCAGGGGGCGCCTGGCCGGGGAGAAGCGCGGGGGCTGGAGCACCA 120
Qy        121 CCAACTGGAGGGTCCGGAGTAGCGAGCGCCCCGAAGGAGGCCATCGGGGAGCCGGGAGGG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CCAACTGGAGGGTCCGGAGTAGCGAGCGCCCCGAAGGAGGCCATCGGGGAGCCGGGAGGG 180
Qy        181 GGGACTGCGAGAGGACCCCGGCGTCCGGGCTCCCGGTGCCAGCGCTATGAGGCCACTCCT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGGACTGCGAGAGGACCCCGGCGTCCGGGCTCCCGGTGCCAGCGCTATGAGGCCACTCCT 240
Qy        241 CGTCCTGCTGCTCCTGGGCCTGGCGGCCGGCTCGCCCCCACTGGACGACAACAAGATCCC 300
```

```
Db        241 CGTCCTGCTGCTCCTGGGCCTGGCGGCCGGCTCGCCCCCACTGGACGACAACAAGATCCC 300
Qy        301 CAGCCTCTGCCCGGGGCACCCCGGCCTTCCAGGCACGCCGGGCCACCATGGCAGCCAGGG 360
Db        301 CAGCCTCTGCCCGGGGCACCCCGGCCTTCCAGGCACGCCGGGCCACCATGGCAGCCAGGG 360
Qy        361 CTTGCCGGGCCGCGATGGCCGCGACGGCCGCGACGGCGCGCCCGGGGCTCCGGGAGAGAA 420
Db        361 CTTGCCGGGCCGCGATGGCCGCGACGGCCGCGACGGCGCGCCCGGGGCTCCGGGAGAGAA 420
Qy        421 AGGCGAGGGCGGGAGGCCGGGACTGCCGGGACCTCGAGGGGACCCCGGGCCGCGAGGAGA 480
Db        421 AGGCGAGGGCGGGAGGCCGGGACTGCCGGGACCTCGAGGGGACCCCGGGCCGCGAGGAGA 480
Qy        481 GGCGGGACCCGCGGGGCCCACCGGGCCTGCCGGGGAGTGCTCGGTGCCTCCGCGATCCGC 540
Db        481 GGCGGGACCCGCGGGGCCCACCGGGCCTGCCGGGGAGTGCTCGGTGCCTCCGCGATCCGC 540
Qy        541 CTTCAGCGCCAAGCGCTCCGAGAGCCGGGTGCCTCCGCCGTCTGACGCACCCTTGCCCTT 600
Db        541 CTTCAGCGCCAAGCGCTCCGAGAGCCGGGTGCCTCCGCCGTCTGACGCACCCTTGCCCTT 600
Qy        601 CGACCGCGTGCTGGTGAACGAGCAGGGACATTACGACGCCGTCACCGGCAAGTTCACCTG 660
Db        601 CGACCGCGTGCTGGTGAACGAGCAGGGACATTACGACGCCGTCACCGGCAAGTTCACCTG 660
Qy        661 CCAGGTGCCTGGGGTCTACTACTTCGCCGTCCATGCCACCGTCTACCGGGCCAGCCTGCA 720
Db        661 CCAGGTGCCTGGGGTCTACTACTTCGCCGTCCATGCCACCGTCTACCGGGCCAGCCTGCA 720
Qy        721 GTTTGATCTGGTGAAGAATGGCGAATCCATTGCCTCTTTCTTCCAGTTTTTCGGGGGGTG 780
Db        721 GTTTGATCTGGTGAAGAATGGCGAATCCATTGCCTCTTTCTTCCAGTTTTTCGGGGGGTG 780
Qy        781 GCCCAAGCCAGCCTCGCTCTCGGGGGGGGCCATGGTGAGGCTGGAGCCTGAGGACCAAGT 840
Db        781 GCCCAAGCCAGCCTCGCTCTCGGGGGGGGCCATGGTGAGGCTGGAGCCTGAGGACCAAGT 840
Qy        841 GTGGGTGCAGGTGGGTGTGGGTGACTACATTGGCATCTATGCCAGCATCAAGACAGACAG 900
Db        841 GTGGGTGCAGGTGGGTGTGGGTGACTACATTGGCATCTATGCCAGCATCAAGACAGACAG 900
Qy        901 CACCTTCTCCGGATTTCTGGTGTACTCCGACTGGCACAGCTCCCCAGTCTTTGCTTAGTG 960
Db        901 CACCTTCTCCGGATTTCTGGTGTACTCCGACTGGCACAGCTCCCCAGTCTTTGCTTAGTG 960
Qy        961 CCCACTGCAAAGTGAGCTCATGCTCTCACTCCTAGAAGGAGGGTGTGAGGCTGACAACCA 1020
Db        961 CCCACTGCAAAGTGAGCTCATGCTCTCACTCCTAGAAGGAGGGTGTGAGGCTGACAACCA 1020
Qy       1021 GGTCATCCAGGAGGGCTGGCCCCCCTGGAATATTGTGAATGACTAGGGAGGTGGGGTAGA 1080
Db       1021 GGTCATCCAGGAGGGCTGGCCCCCCTGGAATATTGTGAATGACTAGGGAGGTGGGGTAGA 1080
Qy       1081 GCACTCTCCGTCCTGCTGCTGGCAAGGAATGGGAACAGTGGCTGTCTGCGATCAGGTCTG 1140
Db       1081 GCACTCTCCGTCCTGCTGCTGGCAAGGAATGGGAACAGTGGCTGTCTGCGATCAGGTCTG 1140
Qy       1141 GCAGCATGGGGCAGTGGCTGGATTTCTGCCCAAGACCAGAGGAGTGTGCTGTGCTGGCAA 1200
Db       1141 GCAGCATGGGGCAGTGGCTGGATTTCTGCCCAAGACCAGAGGAGTGTGCTGTGCTGGCAA 1200
Qy       1201 GTGTAAGTCCCCCAGTTGCTCTGGTCCAGGAGCCCACGGTGGGGTGCTCTCTTCCTGGTC 1260
Db       1201 GTGTAAGTCCCCCAGTTGCTCTGGTCCAGGAGCCCACGGTGGGGTGCTCTCTTCCTGGTC 1260
Qy       1261 CTCTGCTTCTCTGGATCCTCCCCACCCCCTCCTGCTCCTGGGGCCGGCCCTTTTCTCAGA 1320
Db       1261 CTCTGCTTCTCTGGATCCTCCCCACCCCCTCCTGCTCCTGGGGCCGGCCCTTTTCTCAGA 1320
Qy       1321 GATCACTCAATAAACCTAAGAACCCTCATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1377
```

```
Db       1321 GATCACTCAATAAACCTAAGAACCCTCATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1377

RESULT 7
AAC58626
ID   AAC58626 standard; cDNA; 1377 BP.
XX
AC   AAC58626;
XX
DT   29-JAN-2001  (first entry)
XX
DE   Human PRO344 protein UNQ303 encoding cDNA SEQ ID NO:240.
XX
KW   Human; immune related disease; diagnosis; antiinflammatory; cardiant;
KW   dermatological; antiarthritic; antirheumatic; immunosuppressive;
KW   haemostatic; antithyroid; antidiabetic; nootropic; neuroprotective;
KW   antianaemic; hepatotropic; virucide; antipsoriatic; antiallergic;
KW   antiasthmatic; systemic lupus erythematosus; rheumatoid arthritis;
KW   osteoarthritis; spondyloarthropathy; systemic sclerosis; sarcoidosis;
KW   idiopathic inflammatory myopathy; Sjogren's syndrome; thyroiditis;
KW   systemic vasculitis; autoimmune haemolytic anaemia; diabetes mellitus;
KW   autoimmune thrombocytopaenia; immune-mediated renal disease;
KW   demyelinating disease; hepatobiliary disease; Whipple's disease;
KW   inflammatory bowel disease; gluten-sensitive enteropathy;
KW   autoimmune disease; immune-mediated skin disease; allergic disease;
KW   immunological disease; transplantation associated disease;
KW   graft rejection; graft-versus-host-disease; ss.
XX
OS   Homo sapiens.
XX
PN   WO200053758-A2.
XX
PD   14-SEP-2000.
XX
PF   02-MAR-2000; 2000WO-US05841.
XX
PR   08-MAR-1999;   99WO-US05028.
PR   10-MAR-1999;   99US-0123618.
PR   12-MAR-1999;   99US-0123957.
PR   23-MAR-1999;   99US-0125775.
PR   12-APR-1999;   99US-0128849.
PR   20-APR-1999;   99WO-US08615.
PR   28-APR-1999;   99US-0131445.
PR   04-MAY-1999;   99US-0132371.
PR   14-MAY-1999;   99US-0134287.
PR   02-JUN-1999;   99WO-US12252.
PR   23-JUN-1999;   99US-0141037.
PR   20-JUL-1999;   99US-0144758.
PR   26-JUL-1999;   99US-0145698.
PR   28-JUL-1999;   99US-0146222.
PR   01-SEP-1999;   99WO-US20111.
PR   08-SEP-1999;   99WO-US20594.
PR   13-SEP-1999;   99WO-US20944.
PR   15-SEP-1999;   99WO-US21090.
PR   15-SEP-1999;   99WO-US21547.
PR   05-OCT-1999;   99WO-US23089.
PR   29-OCT-1999;   99US-0162506.
PR   29-NOV-1999;   99WO-US28214.
PR   30-NOV-1999;   99WO-US28313.
PR   30-NOV-1999;   99WO-US28409.
PR   01-DEC-1999;   99WO-US28301.
PR   01-DEC-1999;   99WO-US28634.
PR   02-DEC-1999;   99WO-US28551.
PR   02-DEC-1999;   99WO-US28564.
PR   02-DEC-1999;   99WO-US28565.
PR   16-DEC-1999;   99WO-US30095.
PR   20-DEC-1999;   99WO-US30999.
PR   30-DEC-1999;   99WO-US31274.
PR   05-JAN-2000; 2000WO-US00219.
PR   06-JAN-2000; 2000WO-US00277.
PR   06-JAN-2000; 2000WO-US00376.
PR   11-FEB-2000; 2000WO-US03565.
PR   18-FEB-2000; 2000WO-US04341.
```

```
PR   18-FEB-2000; 2000WO-US04342.
PR   22-FEB-2000; 2000WO-US04414.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Ashkenazi AJ,  Baker KP,  Goddard A,  Gurney AL,  Hebert C,  Henzel W;
PI   Kabakoff RC,  Lu Y,  Pan J,  Pennica D,  Shelton DL,  Smith V;
PI   Stewart TA,  Tumas D,  Watanabe CK,  Wood WI,  Yan M;
XX
DR   WPI; 2000-572271/53.
DR   P-PSDB; AAB33461.
XX
PT   Sixty four PRO polypeptides, useful in the diagnosis and treatment of
PT   immune related disorders, e.g. systemic lupus erythematosis, rheumatoid
PT   arthritis, osteoarthritis, thyroiditis and diabetes mellitus -
XX
PS   Claim 23; Fig 95; 309pp; English.
XX
CC   The present invention describes sixty four human PRO proteins which can
CC   be used in the treatment of immune related diseases. The human PRO
CC   proteins, anti-PRO antibodies, agonists and antagonists are useful for
CC   treating and diagnosing immune related disorders. The disorders are
CC   selected from systemic lupus erythematosus, rheumatoid arthritis,
CC   osteoarthritis, juvenile chronic arthritis, spondyloarthropathies,
CC   systemic sclerosis, idiopathic inflammatory myopathies, Sjogren's
CC   syndrome, systemic vasculitis, sarcoidosis, autoimmune haemolytic
CC   anaemia, autoimmune thrombocytopaenia, thyroiditis, diabetes mellitus,
CC   immune-mediated renal disease, demyelinating diseases of the central
CC   and peripheral nervous systems, hepatobiliary diseases, inflammatory
CC   bowel disease, gluten-sensitive enteropathy and Whipple's disease,
CC   autoimmune or immune-mediated skin diseases, allergic diseases,
CC   immunological diseases of the lung, and transplantation associated
CC   diseases including graft rejection and graft-versus-host-disease.
CC   AAC58397 to AAC58578 represent PCR primers and hybridisation probes used
CC   in the isolation of human PRO sequences. AAC58579 to AAC58642 and
CC   AAB33414 to AAB33477 represent human PRO polynucleotide and protein
CC   sequences given in the exemplification of the present invention.
XX
SQ   Sequence 1377 BP; 251 A; 423 C; 471 G; 232 T; 0 other;

  Query Match           99.9%;  Score 1375.4;  DB 21;  Length 1377;
  Best Local Similarity  99.9%;  Pred. No. 3.3e-244;
  Matches 1376;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 GACTAGTTCTCTTGGAGTCTGGGAGGAGGAAAGCGGAGCCGGCAGGGAGCGAACCAGGAC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GACTAGTTCTCTTGGAGTCTGGGAGGAGGAAAGCGGAGCCGGCAGGGAGCGAACCAGGAC 60

Qy         61 TGGGGTGACGGCAGGGCAGGGGGCGCCTGGCCGGGGAGAAGCGCGGGGGCTGGAGCACCA 120
              |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db         61 TGGGGTGACGGCAGCGCAGGGGGCGCCTGGCCGGGGAGAAGCGCGGGGGCTGGAGCACCA 120

Qy        121 CCAACTGGAGGGTCCGGAGTAGCGAGCGCCCCGAAGGAGGCCATCGGGGAGCCGGGAGGG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CCAACTGGAGGGTCCGGAGTAGCGAGCGCCCCGAAGGAGGCCATCGGGGAGCCGGGAGGG 180

Qy        181 GGGACTGCGAGAGGACCCCGGCGTCCGGGCTCCCGGTGCCAGCGCTATGAGGCCACTCCT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGGACTGCGAGAGGACCCCGGCGTCCGGGCTCCCGGTGCCAGCGCTATGAGGCCACTCCT 240

Qy        241 CGTCCTGCTGCTCCTGGGCCTGGCGGCCGGCTCGCCCCCACTGGACGACAACAAGATCCC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CGTCCTGCTGCTCCTGGGCCTGGCGGCCGGCTCGCCCCCACTGGACGACAACAAGATCCC 300

Qy        301 CAGCCTCTGCCCGGGGCACCCCGGCCTTCCAGGCACGCCGGGCCACCATGGCAGCCAGGG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CAGCCTCTGCCCGGGGCACCCCGGCCTTCCAGGCACGCCGGGCCACCATGGCAGCCAGGG 360

Qy        361 CTTGCCGGGCCGCGATGGCCGCGACGGCCGCGACGGCGCGCCCGGGGCTCCGGGAGAGAA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CTTGCCGGGCCGCGATGGCCGCGACGGCCGCGACGGCGCGCCCGGGGCTCCGGGAGAGAA 420

Qy        421 AGGCGAGGGCGGGAGGCCGGGACTGCCGGGACCTCGAGGGGACCCCGGGCCGCGAGGAGA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AGGCGAGGGCGGGAGGCCGGGACTGCCGGGACCTCGAGGGGACCCCGGGCCGCGAGGAGA 480

Qy        481 GGCGGGACCCGCGGGGCCCACCGGGCCTGCCGGGGAGTGCTCGGTGCCTCCGCGATCCGC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GGCGGGACCCGCGGGGCCCACCGGGCCTGCCGGGGAGTGCTCGGTGCCTCCGCGATCCGC 540

Qy        541 CTTCAGCGCCAAGCGCTCCGAGAGCCGGGTGCCTCCGCCGTCTGACGCACCCTTGCCCTT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CTTCAGCGCCAAGCGCTCCGAGAGCCGGGTGCCTCCGCCGTCTGACGCACCCTTGCCCTT 600

Qy        601 CGACCGCGTGCTGGTGAACGAGCAGGGACATTACGACGCCGTCACCGGCAAGTTCACCTG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CGACCGCGTGCTGGTGAACGAGCAGGGACATTACGACGCCGTCACCGGCAAGTTCACCTG 660

Qy        661 CCAGGTGCCTGGGGTCTACTACTTCGCCGTCCATGCCACCGTCTACCGGGCCAGCCTGCA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CCAGGTGCCTGGGGTCTACTACTTCGCCGTCCATGCCACCGTCTACCGGGCCAGCCTGCA 720

Qy        721 GTTTGATCTGGTGAAGAATGGCGAATCCATTGCCTCTTTCTTCCAGTTTTTCGGGGGGTG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GTTTGATCTGGTGAAGAATGGCGAATCCATTGCCTCTTTCTTCCAGTTTTTCGGGGGGTG 780

Qy        781 GCCCAAGCCAGCCTCGCTCTCGGGGGGGGCCATGGTGAGGCTGGAGCCTGAGGACCAAGT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GCCCAAGCCAGCCTCGCTCTCGGGGGGGGCCATGGTGAGGCTGGAGCCTGAGGACCAAGT 840

Qy        841 GTGGGTGCAGGTGGGTGTGGGTGACTACATTGGCATCTATGCCAGCATCAAGACAGACAG 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GTGGGTGCAGGTGGGTGTGGGTGACTACATTGGCATCTATGCCAGCATCAAGACAGACAG 900

Qy        901 CACCTTCTCCGGATTTCTGGTGTACTCCGACTGGCACAGCTCCCCAGTCTTTGCTTAGTG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 CACCTTCTCCGGATTTCTGGTGTACTCCGACTGGCACAGCTCCCCAGTCTTTGCTTAGTG 960

Qy        961 CCCACTGCAAAGTGAGCTCATGCTCTCACTCCTAGAAGGAGGGTGTGAGGCTGACAACCA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 CCCACTGCAAAGTGAGCTCATGCTCTCACTCCTAGAAGGAGGGTGTGAGGCTGACAACCA 1020

Qy       1021 GGTCATCCAGGAGGGCTGGCCCCCCTGGAATATTGTGAATGACTAGGGAGGTGGGGTAGA 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GGTCATCCAGGAGGGCTGGCCCCCCTGGAATATTGTGAATGACTAGGGAGGTGGGGTAGA 1080

Qy       1081 GCACTCTCCGTCCTGCTGCTGGCAAGGAATGGGAACAGTGGCTGTCTGCGATCAGGTCTG 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 GCACTCTCCGTCCTGCTGCTGGCAAGGAATGGGAACAGTGGCTGTCTGCGATCAGGTCTG 1140

Qy       1141 GCAGCATGGGGCAGTGGCTGGATTTCTGCCCAAGACCAGAGGAGTGTGCTGTGCTGGCAA 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 GCAGCATGGGGCAGTGGCTGGATTTCTGCCCAAGACCAGAGGAGTGTGCTGTGCTGGCAA 1200

Qy       1201 GTGTAAGTCCCCCAGTTGCTCTGGTCCAGGAGCCCACGGTGGGGTGCTCTCTTCCTGGTC 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 GTGTAAGTCCCCCAGTTGCTCTGGTCCAGGAGCCCACGGTGGGGTGCTCTCTTCCTGGTC 1260

Qy       1261 CTCTGCTTCTCTGGATCCTCCCCACCCCCTCCTGCTCCTGGGGCCGGCCCTTTTCTCAGA 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 CTCTGCTTCTCTGGATCCTCCCCACCCCCTCCTGCTCCTGGGGCCGGCCCTTTTCTCAGA 1320

Qy       1321 GATCACTCAATAAACCTAAGAACCCTCATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1377
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 GATCACTCAATAAACCTAAGAACCCTCATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1377

RESULT 8
AAX24684
ID   AAX24684 standard; cDNA; 1347 BP.
XX
AC   AAX24684;
XX
DT   21-JUN-1999  (first entry)
XX
DE   Human adipocyte-specific protein zsig39 cDNA.
```

```
XX
KW   Adipocyte-specific protein; zsig39; human; fatty acid metabolism;
KW   energy balance; nutrition; antimicrobial; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             198..929
FT                   /*tag= a
FT   sig_peptide     198..251
FT                   /*tag= b
FT                   /note= "alternatively nucleotides 198-242"
FT   mat_peptide     252..926
FT                   /*tag= c
FT                   /note= "alternatively nucleotides 243-926"
XX
PN   WO9910492-A1.
XX
PD   04-MAR-1999.
XX
PF   26-AUG-1998;   98WO-US17724.
XX
PR   26-AUG-1997;   97US-0056983.
XX
PA   (ZYMO ) ZYMOGENETICS INC.
XX
PI   Humes JM,  Sheppard PO;
XX
DR   WPI; 1999-204665/17.
DR   P-PSDB; AAW97984.
XX
PT   Zsig39 protein - used to modulate fatty acid metabolism
XX
PS   Claim 16; 109-111; 132pp; English.
XX
CC   This isolated polynucleotide encodes human adipocyte-specific
CC   protein zsig39 (AAW97984).  Claimed polynucleotides comprise
CC   nucleotides 243-962, 252-962, 285-482, 285-485, 285-485, 285-488,
CC   285-491, 285-491, 491-926 of this isolated polynucleotide.
CC   Novel zsig39 polypeptides were initially identified by querying an
CC   EST database for secretory signal sequences characterised by an
CC   upstream methionine start site, a hydrophobic region of approximately
CC   13 amino acids and a cleavage site.  A single EST sequence was
CC   discovered, and the novel polypeptide encoded by the full-length
CC   cDNA allowed the identification of a homologue relationship with
CC   adipocyte complement related protein Acrp30 and adipocyte secreted
CC   protein apM1.  A full-length clone was obtained from a lung tissue
CC   library.  The zsig39 gene maps to the 11q23.3 region.  Expression
CC   vectors, cultured cells and a method of producing zsig39 polypeptide
CC   are claimed, as well as probes and primers (useful in diagnostic
CC   applications), and a method for modulating free fatty acid
CC   metabolism by administering a zsig39 polypeptide.  The zsig39
CC   polypeptide may also be used in organ preservation, for
CC   cryopreservation, for surgical pretreatment to prevent injury due
CC   to ischaemia and/or inflammation, and as an antimicrobial agent,
CC   promoting lysis or phagocytosis of infectious agents.
XX
SQ   Sequence 1347 BP; 235 A; 421 C; 465 G; 226 T; 0 other;

  Query Match          96.0%;  Score 1322.2;  DB 20;  Length 1347;
  Best Local Similarity  99.8%;  Pred. No. 2e-234;
  Matches 1324;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy        40 CGGCAGGGAGCGAACCAGGACTGGGGTGACGGCAGGGCAGGGGGCGCCTGGCCGGGGAGA 99
             ||  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        11 CGAGAGGGAGCGAACCAGGACTGGGGTGACGGCAGGGCAGGGGGCGCCTGGCCGGGGAGA 70

Qy       100 AGCGCGGGGGCTGGAGCACCACCAACTGGAGGGTCCGGAGTAGCGAGCGCCCCGAAGGAG 159
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        71 AGCGCGGGGGCTGGAGCACCACCAACTGGAGGGTCCGGAGTAGCGAGCGCCCCGAAGGAG 130

Qy       160 GCCATCGGGGAGCCGGGAGGGGGGACTGCGAGAGGACCCCGGCGTCCGGGCTCCCGGTGC 219
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db       131 GCCATCGGGGAGCCGGGAGGGGGGACTGCGAGAGGACCCCGGCGTCCGGGCTCCCGGTGC 190

Qy       220 CAGCGCTATGAGGCCACTCCTCGTCCTGCTGCTCCTGGGCCTGGCGGCCGGCTCGCCCCC 279
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       191 CAGCGCTATGAGGCCACTCCTCGTCCTGCTGCTCCTGGGCCTGGCGGCCGGCTCGCCCCC 250

Qy       280 ACTGGACGACAACAAGATCCCCAGCCTCTGCCCGGGGCACCCCGGCCTTCCAGGCACGCC 339
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       251 ACTGGACGACAACAAGATCCCCAGCCTCTGCCCGGGGCACCCCGGCCTTCCAGGCACGCC 310

Qy       340 GGGCCACCATGGCAGCCAGGGCTTGCCGGGCCGCGATGGCCGCGACGGCCGCGACGGCGC 399
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       311 GGGCCACCATGGCAGCCAGGGCTTGCCGGGCCGCGATGGCCGCGACGGCCGCGACGGCGC 370

Qy       400 GCCCGGGGCTCCGGGAGAGAAAGGCGAGGGCGGGAGGCCGGGACTGCCGGGACCTCGAGG 459
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       371 GCCCGGGGCTCCGGGAGAGAAAGGCGAGGGCGGGAGGCCGGGACTGCCGGGACCTCGAGG 430

Qy       460 GGACCCCGGGCCGCGAGGAGAGGCGGGACCCGCGGGGCCCACCGGGCCTGCCGGGGAGTG 519
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       431 GGACCCCGGGCCGCGAGGAGAGGCGGGACCCGCGGGGCCCACCGGGCCTGCCGGGGAGTG 490

Qy       520 CTCGGTGCCTCCGCGATCCGCCTTCAGCGCCAAGCGCTCCGAGAGCCGGGTGCCTCCGCC 579
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       491 CTCGGTGCCTCCGCGATCCGCCTTCAGCGCCAAGCGCTCCGAGAGCCGGGTGCCTCCGCC 550

Qy       580 GTCTGACGCACCCTTGCCCTTCGACCGCGTGCTGGTGAACGAGCAGGGACATTACGACGC 639
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       551 GTCTGACGCACCCTTGCCCTTCGACCGCGTGCTGGTGAACGAGCAGGGACATTACGACGC 610

Qy       640 CGTCACCGGCAAGTTCACCTGCCAGGTGCCTGGGGTCTACTACTTCGCCGTCCATGCCAC 699
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       611 CGTCACCGGCAAGTTCACCTGCCAGGTGCCTGGGGTCTACTACTTCGCCGTCCATGCCAC 670

Qy       700 CGTCTACCGGGCCAGCCTGCAGTTTGATCTGGTGAAGAATGGCGAATCCATTGCCTCTTT 759
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       671 CGTCTACCGGGCCAGCCTGCAGTTTGATCTGGTGAAGAATGGCGAATCCATTGCCTCTTT 730

Qy       760 CTTCCAGTTTTTCGGGGGGTGGCCCAAGCCAGCCTCGCTCTCGGGGGGGGCCATGGTGAG 819
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       731 CTTCCAGTTTTTCGGGGGGTGGCCCAAGCCAGCCTCGCTCTCGGGGGGGGCCATGGTGAG 790

Qy       820 GCTGGAGCCTGAGGACCAAGTGTGGGTGCAGGTGGGTGTGGGTGACTACATTGGCATCTA 879
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       791 GCTGGAGCCTGAGGACCAAGTGTGGGTGCAGGTGGGTGTGGGTGACTACATTGGCATCTA 850

Qy       880 TGCCAGCATCAAGACAGACAGCACCTTCTCCGGATTTCTGGTGTACTCCGACTGGCACAG 939
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       851 TGCCAGCATCAAGACAGACAGCACCTTCTCCGGATTTCTGGTGTACTCCGACTGGCACAG 910

Qy       940 CTCCCCAGTCTTTGCTTAGTGCCCACTGCAAAGTGAGCTCATGCTCTCACTCCTAGAAGG 999
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       911 CTCCCCAGTCTTTGCTTAGTGCCCACTGCAAAGTGAGCTCATGCTCTCACTCCTAGAAGG 970

Qy      1000 AGGGTGTGAGGCTGACAACCAGGTCATCCAGGAGGGCTGGCCCCCCTGGAATATTGTGAA 1059
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       971 AGGGTGTGAGGCTGACAACCAGGTCATCCAGGAGGGCTGGCCCCCCTGGAATATTGTGAA 1030

Qy      1060 TGACTAGGGAGGTGGGGTAGAGCACTCTCCGTCCTGCTGCTGGCAAGGAATGGGAACAGT 1119
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1031 TGACTAGGGAGGTGGGGTAGAGCACTCTCCGTCCTGCTGCTGGCAAGGAATGGGAACAGT 1090

Qy      1120 GGCTGTCTGCGATCAGGTCTGGCAGCATGGGGCAGTGGCTGGATTTCTGCCCAAGACCAG 1179
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1091 GGCTGTCTGCGATCAGGTCTGGCAGCATGGGGCAGTGGCTGGATTTCTGCCCAAGACCAG 1150

Qy      1180 AGGAGTGTGCTGTGCTGGCAAGTGTAAGTCCCCCAGTTGCTCTGGTCCAGGAGCCCACGG 1239
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1151 AGGAGTGTGCTGTGCTGGCAAGTGTAAGTCCCCCAGTTGCTCTGGTCCAGGAGCCCACGG 1210

Qy      1240 TGGGGTGCTCTCTTCCTGGTCCTCTGCTTCTCTGGATCCTCCCCACCCCCTCCTGCTCCT 1299
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1211 TGGGGTGCTCTCTTCCTGGTCCTCTGCTTCTCTGGATCCTCCCCACCCCCTCCTGCTCCT 1270
```

```
Qy       1300 GGGGCCGGCCCTTTTCTCAGAGATCACTCAATAAACCTAAGAACCCTCATAAAAAAAAAA 1359
             ||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db       1271 GGGGCCGGCCCTTTTCTCAGAGATCACTCAATAAACCTAAGAACCCTCAAAAAAAAAAAA 1330

Qy       1360 AAAAAAA 1366
              |||||||
Db       1331 AAAAAAA 1337
```

RESULT 9
AAF44970
ID  AAF44970 standard; cDNA; 1338 BP.
XX
AC  AAF44970;
XX
DT  28-MAR-2001  (first entry)
XX
DE  Human TANGO 253 coding sequence SEQ ID NO: 1.
XX
KW  Human; mouse; secreted protein; TANGO253; TANGO 257; TANGO 281;
KW  INTERCEPT 258; coronary disorder; olfactory disorder;
KW  neurological disorder; pulmonary disorder; immunological disorder;
KW  developmental disorder; kidney disorder; ss.
XX
OS  Homo sapiens.
XX
PN  WO200078808-A1.
XX
PD  28-DEC-2000.
XX
PF  19-JUN-2000; 2000WO-US16883.
XX
PR  18-JUN-1999;  99US-0336536.
XX
PA  (MILL-) MILLENNIUM PHARM INC.
XX
PI  Leiby KR,  McKay C,  Bossone S;
XX
DR  WPI; 2001-050109/06.
XX
PT  New nucleic acids for treating diseases and disorders, e.g.
PT  atherosclerosis, infection, autoimmune diseases, obesity, ear
PT  disorders, brain disorders, tumors, diabetes, arthritis, multiple
PT  sclerosis and asthma.
XX
PS  Claim 1; Page 210; 332pp; English.
XX
CC  The present invention provides the protein and coding sequences of the
CC  human and murine secreted or transmembrane proteins TANGO 253, TANGO 257,
CC  TANGO 281 and INTERCEPT 258. These are useful in the treatment of
CC  coronary, pulmonary, olfactory, immunological, neurological,
CC  developmental and kidney disorders.
XX
SQ  Sequence 1338 BP; 229 A; 423 C; 460 G; 226 T; 0 other;

Query Match  94.2%;  Score 1296.6;  DB 22;  Length 1338;
Best Local Similarity  99.6%;  Pred. No. 1e-229;
Matches 1310;  Conservative  0;  Mismatches  4;  Indels  1;  Gaps  1;

```
Qy         54 CCAGGACTGGGGTGACGGCAGGGCAGGGGGCGCCTGGCCGGGGAGAAGCGCGGGGGCTGG 113
              || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         15 CCGGGACTGGGGTGACGGCAGGGCAGGGGGCGCCTGGCCGGGGAGAAGCGCGGGGGCTGG 74

Qy        114 AGCACCACCAACTGGAGGGTCCGGAGTAGCGAGCGCCCCGAAGGAGGCCATCGGGGAGCC 173
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         75 AGCACCACCAACTGGAGGGTCCGGAGTAGCGAGCGCCCCGAAGGAGGCCATCGGGGAGCC 134

Qy        174 GGGAGGGGGGACTGCGAGAGGACCCCGGCGTCCGGGCTCCCGGTGCCAGCGCTATGAGGC 233
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        135 GGGAGGGGGGACTGCGAGAGGACCCCGGCGTCCGGGCTCCCGGTGCCAGCGCTATGAGGC 194

Qy        234 CACTCCTCGTCCTGCTGCTCCTGGGCCTGGCGGCCGGCTCGCCCCCACTGGACGACAACA 293
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        195 CACTCCTCGTCCTGCTGCTCCTGGGCCTGGCGGCCGGCTCGCCCCCACTGGACGACAACA 254

Qy        294 AGATCCCCAGCCTCTGCCCGGGGCACCCCGGCCTTCCAGGCACGCCGGGCCACCATGGCA 353
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        255 AGATCCCCAGCCTCTGCCCGGGGCACCCCGGCCTTCCAGGCACGCCGGGCCACCATGGCA 314

Qy        354 GCCAGGGCTTGCCGGGCCGCGATGGCCGCGACGGCCGCGACGGCGCGCCCGGGGCTCCGG 413
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db.       315 GCCAGGGCTTGCCGGGCCGCGATGGCCGCGACGGCCGCGACGGCGCGCCCGGGGCTCCGG 374

Qy        414 GAGAGAAAGGCGAGGGCGGGAGGCCGGGACTGCCGGGACCTCGAGGGGACCCCGGGCCGC 473
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db        375 GAGAGAAAGGCGAGGGCGGGAGG-CGGGACTGCCGGGACCTCGAGGGGACCCCGGGCCGC 433

Qy        474 GAGGAGAGGCGGGACCCGCGGGGCCCACCGGGCCTGCCGGGGAGTGCTCGGTGCCTCCGC 533
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        434 GAGGAGAGGCGGGACCCGCGGGGCCCACCGGGCCTGCCGGGGAGTGCTCGGTGCCTCCGC 493

Qy        534 GATCCGCCTTCAGCGCCAAGCGCTCCGAGAGCCGGGTGCCTCCGCCGTCTGACGCACCCT 593
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        494 GATCCGCCTTCAGCGCCAAGCGCTCCGAGAGCCGGGTGCCTCCGCCGTCTGACGCACCCT 553

Qy        594 TGCCCTTCGACCGCGTGCTGGTGAACGAGCAGGGACATTACGACGCCGTCACCGGCAAGT 653
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        554 TGCCCTTCGACCGCGTGCTGGTGAACGAGCAGGGACATTACGACGCCGTCACCGGCAAGT 613

Qy        654 TCACCTGCCAGGTGCCTGGGGTCTACTACTTCGCCGTCCATGCCACCGTCTACCGGGCCA 713
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        614 TCACCTGCCAGGTGCCTGGGGTCTACTACTTCGCCGTCCATGCCACCGTCTACCGGGCCA 673

Qy        714 GCCTGCAGTTTGATCTGGTGAAGAATGGCGAATCCATTGCCTCTTTCTTCCAGTTTTTCG 773
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        674 GCCTGCAGTTTGATCTGGTGAAGAATGGCGAATCCATTGCCTCTTTCTTCCAGTTTTTCG 733

Qy        774 GGGGGTGGCCCAAGCCAGCCTCGCTCTCGGGGGGGGCCATGGTGAGGCTGGAGCCTGAGG 833
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        734 GGGGGTGGCCCAAGCCAGCCTCGCTCTCGGGGGGGGCCATGGTGAGGCTGGAGCCTGAGG 793

Qy        834 ACCAAGTGTGGGTGCAGGTGGGTGTGGGTGACTACATTGGCATCTATGCCAGCATCAAGA 893
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        794 ACCAAGTGTGGGTGCAGGTGGGTGTGGGTGACTACATTGGCATCTATGCCAGCATCAAGA 853

Qy        894 CAGACAGCACCTTCTCCGGATTTCTGGTGTACTCCGACTGGCACAGCTCCCCAGTCTTTG 953
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        854 CAGACAGCACCTTCTCCGGATTTCTGGTGTACTCCGACTGGCACAGCTCCCCAGTCTTTG 913

Qy        954 CTTAGTGCCCACTGCAAAGTGAGCTCATGCTCTCACTCCTAGAAGGAGGGTGTGAGGCTG 1013
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        914 CTTAGTGCCCACTGCAAAGTGAGCTCATGCTCTCACTCCTAGAAGGAGGGTGTGAGGCTG 973

Qy       1014 ACAACCAGGTCATCCAGGAGGGCTGGCCCCCCCTGGAATATTGTGAATGACTAGGGAGGTG 1073
              |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        974 ACAACCTGGTCATCCAGGAGGGCTGGCCCCCCCTGGAATATTGTGAATGACTAGGGAGGTG 1033

Qy       1074 GGGTAGAGCACTCTCCGTCCTGCTGCTGGCAAGGAATGGGAACAGTGGCTGTCTGCGATC 1133
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1034 GGGTAGAGCACTCTCCGTCCTGCTGCTGGCAAGGAATGGGAACAGTGGCTGTCTGCGATC 1093

Qy       1134 AGGTCTGGCAGCATGGGGCAGTGGCTGGATTTCTGCCCAAGACCAGAGGAGTGTGCTGTG 1193
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1094 AGGTCTGGCAGCATGGGGCAGTGGCTGGATTTCTGCCCAAGACCAGAGGAGTGTGCTGTG 1153

Qy       1194 CTGGCAAGTGTAAGTCCCCCAGTTGCTCTGGTCCAGGAGCCCACGGTGGGGTGCTCTCTT 1253
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1154 CTGGCAAGTGTAAGTCCCCCAGTTGCTCTGGTCCAGGAGCCCACGGTGGGGTGCTCTCTT 1213

Qy       1254 CCTGGTCCTCTGCTTCTCTGGATCCTCCCCACCCCCTCCTGCTCCTGGGGCCGGCCCTTT 1313
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1214 CCTGGTCCTCTGCTTCTCTGGATCCTCCCCACCCCCTCCTGCTCCTGGGGCCGGCCCTTT 1273

Qy       1314 TCTCAGAGATCACTCAATAAACCTAAGAACCCTCATAAAAAAAAAAAAAAAAAA 1368
              |||||||||||||||||||||||||||||||||| ||||||||||||||||||
```

```
Db      1274 TCTCAGAGATCACTCAATAAACCTAAGAACCCTCCAAAAAAAAAAAAAAAAAAAA 1328

RESULT 10
AAF44998
ID   AAF44998 standard; cDNA; 1338 BP.
XX
AC   AAF44998;
XX
DT   28-MAR-2001  (first entry)
XX
DE   Human secreted protein related coding sequence SEQ ID NO: 103.
XX
KW   Human; mouse; secreted protein; TANGO253; TANGO 257; TANGO 281;
KW   INTERCEPT 258; coronary disorder; olfactory disorder;
KW   neurological disorder; pulmonary disorder; immunological disorder;
KW   developmental disorder; kidney disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200078808-A1.
XX
PD   28-DEC-2000.
XX
PF   19-JUN-2000; 2000WO-US16883.
XX
PR   18-JUN-1999;   99US-0336536.
XX
PA   (MILL-) MILLENNIUM PHARM INC.
XX
PI   Leiby KR,  McKay C,  Bossone S;
XX
DR   WPI; 2001-050109/06.
XX
PT   New nucleic acids for treating diseases and disorders, e.g.
PT   atherosclerosis, infection, autoimmune diseases, obesity, ear
PT   disorders, brain disorders, tumors, diabetes, arthritis, multiple
PT   sclerosis and asthma -
XX
PS   Disclosure; Page 271; 332pp; English.
XX
CC   The present invention provides the protein and coding sequences of the
CC   human and murine secreted or transmembrane proteins TANGO 253, TANGO 257,
CC   TANGO 281 and INTERCEPT 258. These are useful in the treatment of
CC   coronary, pulmonary, olfactory, immunological, neurological,
CC   developmental and kidney disorders.
XX
SQ   Sequence 1338 BP; 229 A; 422 C; 460 G; 227 T; 0 other;

  Query Match           94.0%;  Score 1295;  DB 22;  Length 1338;
  Best Local Similarity  99.5%;  Pred. No. 2e-229;
  Matches 1309;  Conservative    0;  Mismatches    5;  Indels    1;  Gaps    1;

Qy        54 CCAGGACTGGGGTGACGGCAGGGCAGGGGGCGCCTGGCCGGGGAGAAGCGCGGGGGCTGG 113
Db        15 CCGGGACTGGGGTGACGGCAGGGCAGGGGGCGCCTGGCCGGGGAGAAGCGCGGGGGCTGG 74

Qy       114 AGCACCACCAACTGGAGGGTCCGGAGTAGCGAGCGCCCCGAAGGAGGCCATCGGGGAGCC 173
Db        75 AGCACCACCAACTGGAGGGTCCGGAGTAGCGAGCGCCCCGAAGGAGGCCATCGGGGAGCC 134

Qy       174 GGGAGGGGGGACTGCGAGAGGACCCCGGCGTCCGGGCTCCCGGTGCCAGCGCTATGAGGC 233
Db       135 GGGAGGGGGGACTGCGAGAGGACCCCGGCGTCCGGGCTCCCGGTGCCAGCGCTATGAGGC 194

Qy       234 CACTCCTCGTCCTGCTGCTCCTGGGCCTGGCGGCCGGCTCGCCCCCACTGGACGACAACA 293
Db       195 CACTCCTCGTCCTGCTGCTCCTGGGCCTGGCGGCCGGCTCGCCCCCACTGGACGACAACA 254

Qy       294 AGATCCCCAGCCTCTGCCCGGGGCACCCCGGCCTTCCAGGCACGCCGGGCCACCATGGCA 353
Db       255 AGATCCCCAGCCTCTGCCCGGGGCACCCCGGCCTTCCAGGCACGCCGGGCCACCATGGCA 314

Qy       354 GCCAGGGCTTGCCGGGCCGCGATGGCCGCGACGGCCGCGACGGCGCGCCCGGGGCTCCGG 413
Db       315 GCCAGGGCTTGCCGGGCCGCGATGGCCGCGACGGCCGCGACGGTGCGCCCGGGGCTCCGG 374

Qy       414 GAGAGAAAGGCGAGGGCGGGAGGCCGGGACTGCCGGGACCTCGAGGGGACCCCGGGCCGC 473
Db       375 GAGAGAAAGGCGAGGGCGGGAGG-CGGGACTGCCGGGACCTCGAGGGGACCCCGGGCCGC 433

Qy       474 GAGGAGAGGCGGGACCCGCGGGGCCCACCGGGCCTGCCGGGGAGTGCTCGGTGCCTCCGC 533
Db       434 GAGGAGAGGCGGGACCCGCGGGGCCCACCGGGCCTGCCGGGGAGTGCTCGGTGCCTCCGC 493

Qy       534 GATCCGCCTTCAGCGCCAAGCGCTCCGAGAGCCGGGTGCCTCCGCCGTCTGACGCACCCT 593
Db       494 GATCCGCCTTCAGCGCCAAGCGCTCCGAGAGCCGGGTGCCTCCGCCGTCTGACGCACCCT 553

Qy       594 TGCCCTTCGACCGCGTGCTGGTGAACGAGCAGGGACATTACGACGCCGTCACCGGCAAGT 653
Db       554 TGCCCTTCGACCGCGTGCTGGTGAACGAGCAGGGACATTACGACGCCGTCACCGGCAAGT 613

Qy       654 TCACCTGCCAGGTGCCTGGGGTCTACTACTTCGCCGTCCATGCCACCGTCTACCGGGCCA 713
Db       614 TCACCTGCCAGGTGCCTGGGGTCTACTACTTCGCCGTCCATGCCACCGTCTACCGGGCCA 673

Qy       714 GCCTGCAGTTTGATCTGGTGAAGAATGGCGAATCCATTGCCTCTTTCTTCCAGTTTTTCG 773
Db       674 GCCTGCAGTTTGATCTGGTGAAGAATGGCGAATCCATTGCCTCTTTCTTCCAGTTTTTCG 733

Qy       774 GGGGGTGGCCCAAGCCAGCCTCGCTCTCGGGGGGGGCCATGGTGAGGCTGGAGCCTGAGG 833
Db       734 GGGGGTGGCCCAAGCCAGCCTCGCTCTCGGGGGGGGCCATGGTGAGGCTGGAGCCTGAGG 793

Qy       834 ACCAAGTGTGGGTGCAGGTGGGTGTGGGTGACTACATTGGCATCTATGCCAGCATCAAGA 893
Db       794 ACCAAGTGTGGGTGCAGGTGGGTGTGGGTGACTACATTGGCATCTATGCCAGCATCAAGA 853

Qy       894 CAGACAGCACCTTCTCCGGATTTCTGGTGTACTCCGACTGGCACAGCTCCCCAGTCTTTG 953
Db       854 CAGACAGCACCTTCTCCGGATTTCTGGTGTACTCCGACTGGCACAGCTCCCCAGTCTTTG 913

Qy       954 CTTAGTGCCCACTGCAAAGTGAGCTCATGCTCTCACTCCTAGAAGGAGGGTGTGAGGCTG 1013
Db       914 CTTAGTGCCCACTGCAAAGTGAGCTCATGCTCTCACTCCTAGAAGGAGGGTGTGAGGCTG 973

Qy      1014 ACAACCAGGTCATCCAGGAGGGCTGGCCCCCCTGGAATATTGTGAATGACTAGGGAGGTG 1073
Db       974 ACAACCTGGTCATCCAGGAGGGCTGGCCCCCCTGGAATATTGTGAATGACTAGGGAGGTG 1033

Qy      1074 GGGTAGAGCACTCTCCGTCCTGCTGCTGGCAAGGAATGGGAACAGTGGCTGTCTGCGATC 1133
Db      1034 GGGTAGAGCACTCTCCGTCCTGCTGCTGGCAAGGAATGGGAACAGTGGCTGTCTGCGATC 1093

Qy      1134 AGGTCTGGCAGCATGGGGCAGTGGCTGGATTTCTGCCCAAGACCAGAGGAGTGTGCTGTG 1193
Db      1094 AGGTCTGGCAGCATGGGGCAGTGGCTGGATTTCTGCCCAAGACCAGAGGAGTGTGCTGTG 1153

Qy      1194 CTGGCAAGTGTAAGTCCCCCAGTTGCTCTGGTCCAGGAGCCCACGGTGGGGTGCTCTCTT 1253
Db      1154 CTGGCAAGTGTAAGTCCCCCAGTTGCTCTGGTCCAGGAGCCCACGGTGGGGTGCTCTCTT 1213

Qy      1254 CCTGGTCCTCTGCTTCTCTGGATCCTCCCCACCCCCTCCTGCTCCTGGGGCCGGCCCTTT 1313
Db      1214 CCTGGTCCTCTGCTTCTCTGGATCCTCCCCACCCCCTCCTGCTCCTGGGGCCGGCCCTTT 1273

Qy      1314 TCTCAGAGATCACTCAATAAACCTAAGAACCCTCATAAAAAAAAAAAAAAAAAAAA 1368
Db      1274 TCTCAGAGATCACTCAATAAACCTAAGAACCCTCCAAAAAAAAAAAAAAAAAAAA 1328

RESULT 11
AAF44999
ID   AAF44999 standard; cDNA; 1338 BP.
XX
```

```
AC   AAF44999;
XX
DT   28-MAR-2001  (first entry)
XX
DE   Human secreted protein related coding sequence SEQ ID NO: 105.
XX
KW   Human; mouse; secreted protein; TANGO253; TANGO 257; TANGO 281;
KW   INTERCEPT 258; coronary disorder; olfactory disorder;
KW   neurological disorder; pulmonary disorder; immunological disorder;
KW   developmental disorder; kidney disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200078808-A1.
XX
PD   28-DEC-2000.
XX
PF   19-JUN-2000; 2000WO-US16883.
XX
PR   18-JUN-1999;   99US-0336536.
XX
PA   (MILL-) MILLENNIUM PHARM INC.
XX
PI   Leiby KR,  McKay C,  Bossone S;
XX
DR   WPI; 2001-050109/06.
XX
PT   New nucleic acids for treating diseases and disorders, e.g.
PT   atherosclerosis, infection, autoimmune diseases, obesity, ear
PT   disorders, brain disorders, tumors, diabetes, arthritis, multiple
PT   sclerosis and asthma -
XX
PS   Disclosure; Page 272; 332pp; English.
XX
CC   The present invention provides the protein and coding sequences of the
CC   human and murine secreted or transmembrane proteins TANGO 253, TANGO 257,
CC   TANGO 281 and INTERCEPT 258. These are useful in the treatment of
CC   coronary, pulmonary, olfactory, immunological, neurological,
CC   developmental and kidney disorders.
XX
SQ   Sequence 1338 BP; 229 A; 422 C; 460 G; 227 T; 0 other;

  Query Match             94.0%;  Score 1295;  DB 22;  Length 1338;
  Best Local Similarity   99.5%;  Pred. No. 2e-229;
  Matches 1309;  Conservative    0;  Mismatches    5;  Indels    1;  Gaps    1;

Qy         54 CCAGGACTGGGGTGACGGCAGGGCAGGGGGCGCCTGGCCGGGGAGAAGCGCGGGGGCTGG 113
               || |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         15 CCGGGACTGGGGTGACGGCAGGGCAGGGGGCGCCTGGCCGGGGAGAAGCGCGGGGGCTGG  74

Qy        114 AGCACCACCAACTGGAGGGTCCGGAGTAGCGAGCGCCCCGAAGGAGGCCATCGGGGAGCC 173
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         75 AGCACCACCAACTGGAGGGTCCGGAGTAGCGAGCGCCCCGAAGGAGGCCATCGGGGAGCC 134

Qy        174 GGGAGGGGGGACTGCGAGAGGACCCCGGCGTCCGGGCTCCCGGTGCCAGCGCTATGAGGC 233
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        135 GGGAGGGGGGACTGCGAGAGGACCCCGGCGTCCGGGCTCCCGGTGCCAGCGCTATGAGGC 194

Qy        234 CACTCCTCGTCCTGCTGCTCCTGGGCCTGGCGGCCGGCTCGCCCCCACTGGACGACAACA 293
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        195 CACTCCTCGTCCTGCTGCTCCTGGGCCTGGCGGCCGGCTCGCCCCCACTGGACGACAACA 254

Qy        294 AGATCCCCAGCCTCTGCCCGGGGCACCCCGGCCTTCCAGGCACGCCGGGCCACCATGGCA 353
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        255 AGATCCCCAGCCTCTGCCCGGGGCACCCCGGCCTTCCAGGCACGCCGGGCCACCATGGCA 314

Qy        354 GCCAGGGCTTGCCGGGCCGCGATGGCCGCGACGGCCGCGACGGCGCGCCCGGGGCTCCGG 413
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        315 GCCAGGGCTTGCCGGGCCGCGATGGCCGCGACGGCCGCGACGGCGCGCCCGGGGCTCCGG 374

Qy        414 GAGAGAAAGGCGAGGGCGGGAGGCCGGGACTGCCGGGACCTCGAGGGGACCCCGGGCCGC 473
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db        375 GAGAGAAAGGCGAGGGCGGGAGG-CGGGACTGCCGGGACCTCGAGGGGACCCCGGGCCGC 433

Qy        474 GAGGAGAGGCGGGACCCGCGGGGCCCACCGGGCCTGCCGGGGAGTGCTCGGTGCCTCCGC 533
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        434 GAGGAGAGGCGGGACCCGCGGGGCCCACCGGGCCTGTCGGGGAGTGCTCGGTGCCTCCGC 493

Qy        534 GATCCGCCTTCAGCGCCAAGCGCTCCGAGAGCCGGGTGCCTCCGCCGTCTGACGCACCCT 593
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        494 GATCCGCCTTCAGCGCCAAGCGCTCCGAGAGCCGGGTGCCTCCGCCGTCTGACGCACCCT 553

Qy        594 TGCCCTTCGACCGCGTGCTGGTGAACGAGCAGGGACATTACGACGCCGTCACCGGCAAGT 653
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        554 TGCCCTTCGACCGCGTGCTGGTGAACGAGCAGGGACATTACGACGCCGTCACCGGCAAGT 613

Qy        654 TCACCTGCCAGGTGCCTGGGGTCTACTACTTCGCCGTCCATGCCACCGTCTACCGGGCCA 713
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        614 TCACCTGCCAGGTGCCTGGGGTCTACTACTTCGCCGTCCATGCCACCGTCTACCGGGCCA 673

Qy        714 GCCTGCAGTTTGATCTGGTGAAGAATGGCGAATCCATTGCCTCTTTCTTCCAGTTTTTCG 773
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        674 GCCTGCAGTTTGATCTGGTGAAGAATGGCGAATCCATTGCCTCTTTCTTCCAGTTTTTCG 733

Qy        774 GGGGGTGGCCCAAGCCAGCCTCGCTCTCGGGGGGGGCCATGGTGAGGCTGGAGCCTGAGG 833
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        734 GGGGGTGGCCCAAGCCAGCCTCGCTCTCGGGGGGGGCCATGGTGAGGCTGGAGCCTGAGG 793

Qy        834 ACCAAGTGTGGGTGCAGGTGGGTGTGGGTGACTACATTGGCATCTATGCCAGCATCAAGA 893
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        794 ACCAAGTGTGGGTGCAGGTGGGTGTGGGTGACTACATTGGCATCTATGCCAGCATCAAGA 853

Qy        894 CAGACAGCACCTTCTCCGGATTTCTGGTGTACTCCGACTGGCACAGCTCCCCAGTCTTTG 953
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        854 CAGACAGCACCTTCTCCGGATTTCTGGTGTACTCCGACTGGCACAGCTCCCCAGTCTTTG 913

Qy        954 CTTAGTGCCCACTGCAAAGTGAGCTCATGCTCTCACTCCTAGAAGGAGGGTGTGAGGCTG 1013
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        914 CTTAGTGCCCACTGCAAAGTGAGCTCATGCTCTCACTCCTAGAAGGAGGGTGTGAGGCTG 973

Qy       1014 ACAACCAGGTCATCCAGGAGGGCTGGCCCCCCTGGAATATTGTGAATGACTAGGGAGGTG 1073
              |||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        974 ACAACCTGGTCATCCAGGAGGGCTGGCCCCCCTGGAATATTGTGAATGACTAGGGAGGTG 1033

Qy       1074 GGGTAGAGCACTCTCCGTCCTGCTGCTGGCAAGGAATGGGAACAGTGGCTGTCTGCGATC 1133
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1034 GGGTAGAGCACTCTCCGTCCTGCTGCTGGCAAGGAATGGGAACAGTGGCTGTCTGCGATC 1093

Qy       1134 AGGTCTGGCAGCATGGGGCAGTGGCTGGATTTCTGCCCAAGACCAGAGGAGTGTGCTGTG 1193
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1094 AGGTCTGGCAGCATGGGGCAGTGGCTGGATTTCTGCCCAAGACCAGAGGAGTGTGCTGTG 1153

Qy       1194 CTGGCAAGTGTAAGTCCCCCAGTTGCTCTGGTCCAGGAGCCCACGGTGGGGTGCTCTCTT 1253
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1154 CTGGCAAGTGTAAGTCCCCCAGTTGCTCTGGTCCAGGAGCCCACGGTGGGGTGCTCTCTT 1213

Qy       1254 CCTGGTCCTCTGCTTCTCTGGATCCTCCCCACCCCCTCCTGCTCCTGGGGCCGGCCCTTT 1313
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1214 CCTGGTCCTCTGCTTCTCTGGATCCTCCCCACCCCCTCCTGCTCCTGGGGCCGGCCCTTT 1273

Qy       1314 TCTCAGAGATCACTCAATAAACCTAAGAACCCTCATAAAAAAAAAAAAAAAAAAA 1368
              |||||||||||||||||||||||||||||||||||  ||||||||||||||||||
Db       1274 TCTCAGAGATCACTCAATAAACCTAAGAACCCTCCAAAAAAAAAAAAAAAAAAAA 1328

RESULT 12
AAF45000
ID   AAF45000 standard; cDNA; 1338 BP.
XX
AC   AAF45000;
XX
DT   28-MAR-2001  (first entry)
XX
DE   Human secreted protein related coding sequence SEQ ID NO: 107.
XX
KW   Human; mouse; secreted protein; TANGO253; TANGO 257; TANGO 281;
```

```
KW   INTERCEPT 258; coronary disorder; olfactory disorder;
KW   neurological disorder; pulmonary disorder; immunological disorder;
KW   developmental disorder; kidney disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200078808-A1.
XX
PD   28-DEC-2000.
XX
PF   19-JUN-2000; 2000WO-US16883.
XX
PR   18-JUN-1999;   99US-0336536.
XX
PA   (MILL-) MILLENNIUM PHARM INC.
XX
PI   Leiby KR,  McKay C,  Bossone S;
XX
DR   WPI; 2001-050109/06.
XX
PT   New nucleic acids for treating diseases and disorders, e.g.
PT   atherosclerosis, infection, autoimmune diseases, obesity, ear
PT   disorders, brain disorders, tumors, diabetes, arthritis, multiple
PT   sclerosis and asthma -
XX
PS   Disclosure; Page 273; 332pp; English.
XX
CC   The present invention provides the protein and coding sequences of the
CC   human and murine secreted or transmembrane proteins TANGO 253, TANGO 257,
CC   TANGO 281 and INTERCEPT 258. These are useful in the treatment of
CC   coronary, pulmonary, olfactory, immunological, neurological,
CC   developmental and kidney disorders.
XX
SQ   Sequence 1338 BP; 229 A; 424 C; 460 G; 225 T; 0 other;
```

Query Match 94.0%; Score 1295; DB 22; Length 1338;
Best Local Similarity 99.5%; Pred. No. 2e-229;
Matches 1309; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

```
Qy      54 CCAGGACTGGGGTGACGGCAGGGCAGGGGGCGCCTGGCCGGGGAGAAGCGCGGGGGCTGG 113
           || |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      15 CCGGGACTGGGGTGACGGCAGGGCAGGGGGCGCCTGGCCGGGGAGAAGCGCGGGGGCTGG 74

Qy     114 AGCACCACCAACTGGAGGGTCCGGAGTAGCGAGCGCCCCGAAGGAGGCCATCGGGGAGCC 173
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      75 AGCACCACCAACTGGAGGGTCCGGAGTAGCGAGCGCCCCGAAGGAGGCCATCGGGGAGCC 134

Qy     174 GGGAGGGGGGACTGCGAGAGGACCCCGGCGTCCGGGCTCCCGGTGCCAGCGCTATGAGGC 233
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     135 GGGAGGGGGGACTGCGAGAGGACCCCGGCGTCCGGGCTCCCGGTGCCAGCGCTATGAGGC 194

Qy     234 CACTCCTCGTCCTGCTGCTCCTGGGCCTGGCGGCCGGCTCGCCCCCACTGGACGACAACA 293
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     195 CACTCCTCGTCCTGCTGCTCCTGGGCCTGGCGGCCGGCTCGCCCCCACTGGACGACAACA 254

Qy     294 AGATCCCCAGCCTCTGCCCGGGGCACCCCGGCCTTCCAGGCACGCCGGGCCACCATGGCA 353
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     255 AGATCCCCAGCCTCTGCCCGGGGCACCCCGGCCTTCCAGGCACGCCGGGCCACCATGGCA 314

Qy     354 GCCAGGGCTTGCCGGGCCGCGATGGCCGCGACGGCCGCGACGGCGCGCCCGGGGCTCCGG 413
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     315 GCCAGGGCTTGCCGGGCCGCGATGGCCGCGACGGCCGCGACGGCGCGCCCGGGGCTCCGG 374

Qy     414 GAGAGAAAGGCGAGGGCGGGAGGCCGGGACTGCCGGGACCTCGAGGGGACCCCGGGCCGC 473
           |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db     375 GAGAGAAAGGCGAGGGCGGGAGG-CGGGACTGCCGGGACCTCGAGGGGACCCCGGGCCGC 433

Qy     474 GAGGAGAGGCGGGACCCGCGGGGCCCACCGGGCCTGCCGGGGAGTGCTCGGTGCCTCCGC 533
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     434 GAGGAGAGGCGGGACCCGCGGGGCCCACCGGGCCTGCCGGGGAGTGCTCGGTGCCTCCGC 493

Qy     534 GATCCGCCTTCAGCGCCAAGCGCTCCGAGAGCCGGGTGCCTCCGCCGTCTGACGCACCCT 593
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db     494 GATCCGCCTTCAGCGCCAAGCGCTCCGAGAGCCGGGTGCCTCCGCCGTCTGACGCACCCT 553

Qy     594 TGCCCTTCGACCGCGTGCTGGTGAACGAGCAGGGACATTACGACGCCGTCACCGGCAAGT 653
           |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db     554 TGCCCTTCGACCGCGTGCTGGCGAACGAGCAGGGACATTACGACGCCGTCACCGGCAAGT 613

Qy     654 TCACCTGCCAGGTGCCTGGGGTCTACTACTTCGCCGTCCATGCCACCGTCTACCGGGCCA 713
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     614 TCACCTGCCAGGTGCCTGGGGTCTACTACTTCGCCGTCCATGCCACCGTCTACCGGGCCA 673

Qy     714 GCCTGCAGTTTGATCTGGTGAAGAATGGCGAATCCATTGCCTCTTTCTTCCAGTTTTTCG 773
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     674 GCCTGCAGTTTGATCTGGTGAAGAATGGCGAATCCATTGCCTCTTTCTTCCAGTTTTTCG 733

Qy     774 GGGGGTGGCCCAAGCCAGCCTCGCTCTCGGGGGGGGCCATGGTGAGGCTGGAGCCTGAGG 833
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     734 GGGGGTGGCCCAAGCCAGCCTCGCTCTCGGGGGGGGCCATGGTGAGGCTGGAGCCTGAGG 793

Qy     834 ACCAAGTGTGGGTGCAGGTGGGTGTGGGTGACTACATTGGCATCTATGCCAGCATCAAGA 893
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     794 ACCAAGTGTGGGTGCAGGTGGGTGTGGGTGACTACATTGGCATCTATGCCAGCATCAAGA 853

Qy     894 CAGACAGCACCTTCTCCGGATTTCTGGTGTACTCCGACTGGCACAGCTCCCCAGTCTTTG 953
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     854 CAGACAGCACCTTCTCCGGATTTCTGGTGTACTCCGACTGGCACAGCTCCCCAGTCTTTG 913

Qy     954 CTTAGTGCCCACTGCAAAGTGAGCTCATGCTCTCACTCCTAGAAGGAGGGTGTGAGGCTG 1013
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     914 CTTAGTGCCCACTGCAAAGTGAGCTCATGCTCTCACTCCTAGAAGGAGGGTGTGAGGCTG 973

Qy    1014 ACAACCAGGTCATCCAGGAGGGCTGGCCCCCCTGGAATATTGTGAATGACTAGGGAGGTG 1073
           |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     974 ACAACCTGGTCATCCAGGAGGGCTGGCCCCCCTGGAATATTGTGAATGACTAGGGAGGTG 1033

Qy    1074 GGGTAGAGCACTCTCCGTCCTGCTGCTGGCAAGGAATGGGAACAGTGGCTGTCTGCGATC 1133
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1034 GGGTAGAGCACTCTCCGTCCTGCTGCTGGCAAGGAATGGGAACAGTGGCTGTCTGCGATC 1093

Qy    1134 AGGTCTGGCAGCATGGGGCAGTGGCTGGATTTCTGCCCAAGACCAGAGGAGTGTGCTGTG 1193
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1094 AGGTCTGGCAGCATGGGGCAGTGGCTGGATTTCTGCCCAAGACCAGAGGAGTGTGCTGTG 1153

Qy    1194 CTGGCAAGTGTAAGTCCCCCAGTTGCTCTGGTCCAGGAGCCCACGGTGGGGTGCTCTCTT 1253
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1154 CTGGCAAGTGTAAGTCCCCCAGTTGCTCTGGTCCAGGAGCCCACGGTGGGGTGCTCTCTT 1213

Qy    1254 CCTGGTCCTCTGCTTCTCTGGATCCTCCCCACCCCCTCCTGCTCCTGGGGCCGGCCCTTT 1313
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1214 CCTGGTCCTCTGCTTCTCTGGATCCTCCCCACCCCCTCCTGCTCCTGGGGCCGGCCCTTT 1273

Qy    1314 TCTCAGAGATCACTCAATAAACCTAAGAACCCTCATAAAAAAAAAAAAAAAAAAA 1368
           |||||||||||||||||||||||||||||||||  ||||||||||||||||||||
Db    1274 TCTCAGAGATCACTCAATAAACCTAAGAACCCTCCAAAAAAAAAAAAAAAAAAAA 1328
```

```
RESULT 13
AAF45001
ID   AAF45001 standard; cDNA; 1338 BP.
XX
AC   AAF45001;
XX
DT   28-MAR-2001  (first entry)
XX
DE   Human secreted protein related coding sequence SEQ ID NO: 109.
XX
KW   Human; mouse; secreted protein; TANGO253; TANGO 257; TANGO 281;
KW   INTERCEPT 258; coronary disorder; olfactory disorder;
KW   neurological disorder; pulmonary disorder; immunological disorder;
KW   developmental disorder; kidney disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200078808-A1.
```

```
XX
PD   28-DEC-2000.
XX
PF   19-JUN-2000; 2000WO-US16883.
XX
PR   18-JUN-1999;   99US-0336536.
XX
PA   (MILL-) MILLENNIUM PHARM INC.
XX
PI   Leiby KR,  McKay C,  Bossone S;
XX
DR   WPI; 2001-050109/06.
XX
PT   New nucleic acids for treating diseases and disorders; e.g.
PT   atherosclerosis, infection, autoimmune diseases, obesity, ear
PT   disorders, brain disorders, tumors, diabetes, arthritis, multiple
PT   sclerosis and asthma -
XX
PS   Disclosure; Page 274-275; 332pp; English.
XX
CC   The present invention provides the protein and coding sequences of the
CC   human and murine secreted or transmembrane proteins TANGO 253, TANGO 257,
CC   TANGO 281 and INTERCEPT 258. These are useful in the treatment of
CC   coronary, pulmonary, olfactory, immunological, neurological,
CC   developmental and kidney disorders.
XX
SQ   Sequence 1338 BP; 228 A; 424 C; 460 G; 226 T; 0 other;

  Query Match             94.0%;  Score 1295;  DB 22;  Length 1338;
  Best Local Similarity   99.5%;  Pred. No. 2e-229;
  Matches 1309;  Conservative    0;  Mismatches    5;  Indels    1;  Gaps    1;

Qy        54 CCAGGACTGGGGTGACGGCAGGGCAGGGGGCGCCTGGCCGGGGAGAAGCGCGGGGGCTGG 113
Db        15 CCGGGACTGGGGTGACGGCAGGGCAGGGGGCGCCTGGCCGGGGAGAAGCGCGGGGGCTGG  74

Qy       114 AGCACCACCAACTGGAGGGTCCGGAGTAGCGAGCGCCCCGAAGGAGGCCATCGGGGAGCC 173
Db        75 AGCACCACCAACTGGAGGGTCCGGAGTAGCGAGCGCCCCGAAGGAGGCCATCGGGGAGCC 134

Qy       174 GGGAGGGGGGACTGCGAGAGGACCCCGGCGTCCGGGCTCCCGGTGCCAGCGCTATGAGGC 233
Db       135 GGGAGGGGGGACTGCGAGAGGACCCCGGCGTCCGGGCTCCCGGTGCCAGCGCTATGAGGC 194

Qy       234 CACTCCTCGTCCTGCTGCTCCTGGGCCTGGCGGCCGGCTCGCCCCCACTGGACGACAACA 293
Db       195 CACTCCTCGTCCTGCTGCTCCTGGGCCTGGCGGCCGGCTCGCCCCCACTGGACGACAACA 254

Qy       294 AGATCCCCAGCCTCTGCCCGGGGCACCCCGGCCTTCCAGGCACGCCGGGCCACCATGGCA 353
Db       255 AGATCCCCAGCCTCTGCCCGGGGCACCCCGGCCTTCCAGGCACGCCGGGCCACCATGGCA 314

Qy       354 GCCAGGGCTTGCCGGGCCGCGATGGCCGCGACGGCCGCGACGGCGCGCCCCGGGGCTCCGG 413
Db       315 GCCAGGGCTTGCCGGGCCGCGATGGCCGCGACGGCCGCGACGGCGCGCCCCGGGGCTCCGG 374

Qy       414 GAGAGAAAGGCGAGGGCGGGAGGCCGGGACTGCCGGGACCTCGAGGGGACCCCGGGCCGC 473
Db       375 GAGAGAAAGGCGAGGGCGGGAGG-CGGGACTGCCGGGACCTCGAGGGGACCCCGGGCCGC 433

Qy       474 GAGGAGAGGCGGGACCCGCGGGGCCCACCGGGCCTGCCGGGGAGTGCTCGGTGCCTCCGC 533
Db       434 GAGGAGAGGCGGGACCCGCGGGGCCCACCGGGCCTGCCGGGGAGTGCTCGGTGCCTCCGC 493

Qy       534 GATCCGCCTTCAGCGCCAAGCGCTCCGAGAGCCGGGTGCCTCCGCCGTCTGACGCACCCT 593
Db       494 GATCCGCCTTCAGCGCCAAGCGCTCCGAGAGCCGGGTGCCTCCGCCGTCTGACGCACCCT 553

Qy       594 TGCCCTTCGACCGCGTGCTGGTGAACGAGCAGGGACATTACGACGCCGTCACCGGCAAGT 653
Db       554 TGCCCTTCGACCGCGTGCTGGTGAACGAGCAGGGACATTACGACGCCGTCACCGGCAAGT 613

Qy       654 TCACCTGCCAGGTGCCTGGGGTCTACTACTTCGCCGTCCATGCCACCGTCTACCGGGCCA 713
Db       614 TCACCTGCCAGGTGCCTGGGGTCTACTACTTCGCCGTCCATGCCACCGTCTACCGGGCCA 673

Qy       714 GCCTGCAGTTTGATCTGGTGAAGAATGGCGAATCCATTGCCTCTTTCTTCCAGTTTTTCG 773
Db       674 GCCTGCAGTTTGATCTGGTGAAGAATGGCGAATCCCTTGCCTCTTTCTTCCAGTTTTTCG 733

Qy       774 GGGGGTGGCCCAAGCCAGCCTCGCTCTCGGGGGGGGCCATGGTGAGGCTGGAGCCTGAGG 833
Db       734 GGGGGTGGCCCAAGCCAGCCTCGCTCTCGGGGGGGGCCATGGTGAGGCTGGAGCCTGAGG 793

Qy       834 ACCAAGTGTGGGTGCAGGTGGGTGTGGGTGACTACATTGGCATCTATGCCAGCATCAAGA 893
Db       794 ACCAAGTGTGGGTGCAGGTGGGTGTGGGTGACTACATTGGCATCTATGCCAGCATCAAGA 853

Qy       894 CAGACAGCACCTTCTCCGGATTTCTGGTGTACTCCGACTGGCACAGCTCCCCAGTCTTTG 953
Db       854 CAGACAGCACCTTCTCCGGATTTCTGGTGTACTCCGACTGGCACAGCTCCCCAGTCTTTG 913

Qy       954 CTTAGTGCCCACTGCAAAGTGAGCTCATGCTCTCACTCCTAGAAGGAGGGTGTGAGGCTG 1013
Db       914 CTTAGTGCCCACTGCAAAGTGAGCTCATGCTCTCACTCCTAGAAGGAGGGTGTGAGGCTG 973

Qy      1014 ACAACCAGGTCATCCAGGAGGGCTGGCCCCCCTGGAATATTGTGAATGACTAGGGAGGTG 1073
Db       974 ACAACCTGGTCATCCAGGAGGGCTGGCCCCCCTGGAATATTGTGAATGACTAGGGAGGTG 1033

Qy      1074 GGGTAGAGCACTCTCCGTCCTGCTGCTGGCAAGGAATGGGAACAGTGGCTGTCTGCGATC 1133
Db      1034 GGGTAGAGCACTCTCCGTCCTGCTGCTGGCAAGGAATGGGAACAGTGGCTGTCTGCGATC 1093

Qy      1134 AGGTCTGGCAGCATGGGGCAGTGGCTGGATTTCTGCCCAAGACCAGAGGAGTGTGCTGTG 1193
Db      1094 AGGTCTGGCAGCATGGGGCAGTGGCTGGATTTCTGCCCAAGACCAGAGGAGTGTGCTGTG 1153

Qy      1194 CTGGCAAGTGTAAGTCCCCCAGTTGCTCTGGTCCAGGAGCCCACGGTGGGGTGCTCTCTT 1253
Db      1154 CTGGCAAGTGTAAGTCCCCCAGTTGCTCTGGTCCAGGAGCCCACGGTGGGGTGCTCTCTT 1213

Qy      1254 CCTGGTCCTCTCTGCTTCTCTGGATCCTCCCCACCCCCTCCTGCTCCTGGGGCCGGCCCTTT 1313
Db      1214 CCTGGTCCTCTCTGCTTCTCTGGATCCTCCCCACCCCCTCCTGCTCCTGGGGCCGGCCCTTT 1273

Qy      1314 TCTCAGAGATCACTCAATAAACCTAAGAACCCTCATAAAAAAAAAAAAAAAAAAA 1368
Db      1274 TCTCAGAGATCACTCAATAAACCTAAGAACCCTCCAAAAAAAAAAAAAAAAAAA 1328

RESULT 14
AAL39624
ID   AAL39624 standard; DNA; 3248 BP.
XX
AC   AAL39624;
XX
DT   05-SEP-2002  (first entry)
XX
DE   Human secreted protein DNA SEQ ID No 68.
XX
KW   Antiarteriosclerotic; cytostatic; HIV; antiallergic; antianaemic;
KW   antiasthmatic; cardiant; vasotropic; neuroprotective; nootropic; SECP;
KW   anticonvulsant; antiparkinsonian; cerebroprotective; antiinflammatory;
KW   immunosuppressive; human secreted protein; cell proliferative disorder;
KW   arteriosclerosis; cancer; autoimmune; inflammatory disorder; AIDS;
KW   allergy; anaemia; asthma; cardiovascular disease; developmental disorder;
KW   ischaemic heart disease; congestive heart failure; neurological disorder;
KW   renal tubular acidosis; hypothyroidism; Alzheimer's disease; dementia;
KW   Parkinson's disease; epilepsy; stroke; knockin humanised animal;
KW   transgenic animal; gene therapy; gene; ds.
XX
OS   Homo sapiens.
XX
PN   WO200238602-A2.
XX
```

PD 16-MAY-2002.
XX
PF 08-NOV-2001; 2001WO-US47420.
XX
PR 08-NOV-2000; 2000US-247505P.
PR 09-NOV-2000; 2000US-248642P.
PR 16-NOV-2000; 2000US-249824P.
PR 21-NOV-2000; 2000US-252824P.
PR 08-DEC-2000; 2000US-254305P.
PR 18-DEC-2000; 2000US-256448P.
XX
PA (INCY-) INCYTE GENOMICS INC.
XX
PI Yue H, Yao MG, Gandhi AR, Baughn MR, Swarnakar A, Walia NK;
PI Sanjanwala M, Thornton M, Elliott VS, Lu Y, Gietzen KJ, Burford N;
PI Ding L, Hafalia AJA, Tang YT, Bandman O, Warren BA, Honchell CD;
PI Lu DAM, Thangavelu K, Lee S, Xu Y, Yang J, Lal PG, Tran B;
PI Ison CH, Duggan BM, Sapperstein SK;
XX
DR WPI; 2002-519296/55.
DR P-PSDB; AAO21663.
XX
PT Human secreted proteins and polynucleotides for diagnosing, treating or
PT preventing disorders of cell proliferative, cardiovascular,
PT developmental, neurological and autoimmune/inflammatory disorders -
XX
PS Claim 5; Page 194-195; 229pp; English.
XX
CC The invention relates to an isolated human secreted protein (SECP)
CC polypeptide from 63 fully defined protein sequences given in the
CC specification. The polypeptide is useful for the diagnosing/treating of a
CC disease with decreased/overexpression of SECP. Examples of disorders
CC associated with abnormal expression of SECP include a cell proliferative
CC disorder e.g. arteriosclerosis, cancers; autoimmune/inflammatory
CC disorder, AIDS, allergies, anaemia, asthma; cardiovascular disease e.g.
CC congestive heart failure, ischaemic heart disease; developmental disorder
CC e.g. renal tubular acidosis, hypothyroidism; neurological disorder e.g.
CC Alzheimer's disease, dementia, Parkinson's disease, epilepsy or stroke.
CC The SECP polynucleotide and polypeptide are further useful for analysing
CC the proteome of a tissue or a cell type. The polynucleotide is useful for
CC creating knockin humanised animals (pigs) or transgenic animals (mice or
CC rats) to model human disease, and for somatic or germline gene therapy,
CC and further for generating hybridisation probes useful in mapping the
CC naturally occurring genomic sequence. This polynucleotide sequence
CC represents the DNA of a human secreted protein of the invention.
XX
SQ Sequence 3248 BP; 565 A; 1058 C; 946 G; 679 T; 0 other;

Query Match 78.8%; Score 1084.8; DB 24; Length 3248;
Best Local Similarity 90.5%; Pred. No. 1.1e-190;
Matches 1222; Conservative 0; Mismatches 2; Indels 126; Gaps 1;

```
Qy        7 TTCTCTTGGAGTCTGGGAGGAGGAAAGCGGAGCCGGCAGGGAGCGAACCAGGACTGGGGT 66
            | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2025 TCCTCTTGGAGTCTGGGAGGAGGAAAGCGGAGCCGGCAGGGAGCGAACCAGGACTGGGGT 2084

Qy       67 GACGGCAGGGCAGGGGGCGCCTGGCCGGGGAGAAGCGCGGGGGCTGGAGCACCACCAACT 126
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2085 GACGGCAGGGCAGGGGGCGCCTGGCCGGGGAGAAGCGCGGGGGCTGGAGCACCACCAACT 2144

Qy      127 GGAGGGTCCGGAGTAGCGAGCGCCCCGAAGGAGGCCATCGGGGAGCCGGGAGGGGGGACT 186
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2145 GGAGGGTCCGGAGTAGCGAGCGCCCCGAAGGAGGCCATCGGGGAGCCGGGAGGGGGGACT 2204

Qy      187 GCGAGAGGACCCCGGCGTCCGGGCTCCCGGTGCCAGCGCTATGAGGCCACTCCTCGTCCT 246
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2205 GCGAGAGGACCCCGGCGTCCGGGCTCCCGGTGCCAGCGCTATGAGGCCACTCCTCGTCCT 2264

Qy      247 GCTGCTCCTGGGCCTGGCGGCCGGCTCGCCCCCACTGGACGACAACAAGATCCCCAGCCT 306
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2265 GCTGCTCCTGGGCCTGGCGGCCGGCTCGCCCCCACTGGACGACAACAAGATCCCCAGCCT 2324

Qy      307 CTGCCCGGGGCACCCCGGCCTTCCAGGCACGCCGGGCCACCATGGCAGCCAGGGCTTGCC 366
```

```
            ||||
Db     2325 CTGC---------------------------------------------------------- 2328

Qy      367 GGGCCGCGATGGCCGCGACGGCCGCGACGGCGCGCCCGGGGCTCCGGGAGAGAAAGGCGA 426

Db     2329 ---------------------------------------------------------- 2328

Qy      427 GGGCGGGAGGCCGGGACTGCCGGGACCTCGAGGGGACCCCGGGCCGCGAGGAGAGGCGGG 486
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2329 ----------CCGGGACTGCCGGGACCTCGAGGGGACCCCGGGCCGCGAGGAGAGGCGGG 2378

Qy      487 ACCCGCGGGGCCCACCGGGCCTGCCGGGGAGTGCTCGGTGCCTCCGCGATCCGCCTTCAG 546
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2379 ACCCGCGGGGCCCACCGGGCCTGCCGGGGAGTGCTCGGTGCCTCCGCGATCCGCCTTCAG 2438

Qy      547 CGCCAAGCGCTCCGAGAGCCGGGTGCCTCCGCCGTCTGACGCACCCTTGCCCTTCGACCG 606
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2439 CGCCAAGCGCTCCGAGAGCCGGGTGCCTCCGCCGTCTGACGCACCCTTGCCCTTCGACCG 2498

Qy      607 CGTGCTGGTGAACGAGCAGGGACATTACGACGCCGTCACCGGCAAGTTCACCTGCCAGGT 666
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2499 CGTGCTGGTGAACGAGCAGGGACATTACGACGCCGTCACCGGCAAGTTCACCTGCCAGGT 2558

Qy      667 GCCTGGGGTCTACTACTTCGCCGTCCATGCCACCGTCTACCGGGCCAGCCTGCAGTTTGA 726
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2559 GCCTGGGGTCTACTACTTCGCCGTCCATGCCACCGTCTACCGGGCCAGCCTGCAGTTTGA 2618

Qy      727 TCTGGTGAAGAATGGCGAATCCATTGCCTCTTTCTTCCAGTTTTTCGGGGGGTGGCCCAA 786
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2619 TCTGGTGAAGAATGGCGAATCCATTGCCTCTTTCTTCCAGTTTTTCGGGGGGTGGCCCAA 2678

Qy      787 GCCAGCCTCGCTCTCGGGGGGGGCCATGGTGAGGCTGGAGCCTGAGGACCAAGTGTGGGT 846
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2679 GCCAGCCTCGCTCTCGGGGGGGGCCATGGTGAGGCTGGAGCCTGAGGACCAAGTGTGGGT 2738

Qy      847 GCAGGTGGGTGTGGGTGACTACATTGGCATCTATGCCAGCATCAAGACAGACAGCACCTT 906
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2739 GCAGGTGGGTGTGGGTGACTACATTGGCATCTATGCCAGCATCAAGACAGACAGCACCTT 2798

Qy      907 CTCCGGATTTCTGGTGTACTCCGACTGGCACAGCTCCCCAGTCTTTGCTTAGTGCCCACT 966
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2799 CTCCGGATTTCTGGTGTACTCCGACTGGCACAGCTCCCCAGTCTTTGCTTAGTGCCCACT 2858

Qy      967 GCAAAGTGAGCTCATGCTCTCACTCCTAGAAGGAGGGTGTGAGGCTGACAACCAGGTCAT 1026
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2859 GCAAAGTGAGCTCATGCTCTCACTCCTAGAAGGAGGGTGTGAGGCTGACAACCAGGTCAT 2918

Qy     1027 CCAGGAGGGCTGGCCCCCCTGGAATATTGTGAATGACTAGGGAGGTGGGGTAGAGCACTC 1086
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2919 CCAGGAGGGCTGGCCCCCCTGGAATATTGTGAATGACTAGGGAGGTGGGGTAGAGCACTC 2978

Qy     1087 TCCGTCCTGCTGCTGGCAAGGAATGGGAACAGTGGCTGTCTGCGATCAGGTCTGGCAGCA 1146
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2979 TCCGTCCTGCTGCTGGCAAGGAATGGGAACAGTGGCTGTCTGCGATCAGGTCTGGCAGCA 3038

Qy     1147 TGGGGCAGTGGCTGGATTTCTGCCCAAGACCAGAGGAGTGTGCTGTGCTGGCAAGTGTAA 1206
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3039 TGGGGCAGTGGCTGGATTTCTGCCCAAGACCAGAGGAGTGTGCTGTGCTGGCAAGTGTAA 3098

Qy     1207 GTCCCCCAGTTGCTCTGGTCCAGGAGCCCACGGTGGGGTGCTCTCTTCCTGGTCCTCTGC 1266
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3099 GTCCCCCAGTTGCTCTGGTCCAGGAGCCCACGGTGGGGTGCTCTCTTCCTGGTCCTCTGC 3158

Qy     1267 TTCTCTGGATCCTCCCCACCCCCTCCTGCTCCTGGGGCCGGCCCTTTTCTCAGAGATCAC 1326
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3159 TTCTCTGGATCCTCCCCACCCCCTCCTGCTCCTGGGGCCGGCCCTTTTCTCAGAGATCAC 3218

Qy     1327 TCAATAAACCTAAGAACCCTCATAAAAAAA 1356
            ||||||||||||||||||||| |||||||
Db     3219 TCAATAAACCTAAGAACCCTCAAAAAAAAA 3248
```

RESULT 15

```
AAS76911
ID   AAS76911 standard; cDNA; 1082 BP.
XX
AC   AAS76911;
XX
DT   13-FEB-2002  (first entry)
XX
DE   DNA encoding novel human diagnostic protein #12715.
XX
KW   Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW   food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200175067-A2.
XX
PD   11-OCT-2001.
XX
PF   30-MAR-2001; 2001WO-US08631.
XX
PR   31-MAR-2000; 2000US-0540217.
PR   23-AUG-2000; 2000US-0649167.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Drmanac RT,  Liu C,  Tang YT;
XX
DR   WPI; 2001-639362/73.
DR   P-PSDB; ABG12724.
XX
PT   New isolated polynucleotide and encoded polypeptides, useful in
PT   diagnostics, forensics, gene mapping, identification of mutations
PT   responsible for genetic disorders or other traits and to assess
PT   biodiversity -
XX
PS   Claim 1; SEQ ID No 12715; 103pp; English.
XX
CC   The invention relates to isolated polynucleotide (I) and
CC   polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC   polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC   and gene mapping, and in recombinant production of (II). The
CC   polynucleotides are also used in diagnostics as expressed sequence tags
CC   for identifying expressed genes. (I) is useful in gene therapy techniques
CC   to restore normal activity of (II) or to treat disease states involving
CC   (II). (II) is useful for generating antibodies against it, detecting or
CC   quantitating a polypeptide in tissue, as molecular weight markers and as
CC   a food supplement. (II) and its binding partners are useful in medical
CC   imaging of sites expressing (II). (I) and (II) are useful for treating
CC   disorders involving aberrant protein expression or biological activity.
CC   The polypeptide and polynucleotide sequences have applications in
CC   diagnostics, forensics, gene mapping, identification of mutations
CC   responsible for genetic disorders or other traits to assess biodiversity
CC   and to produce other types of data and products dependent on DNA and
CC   amino acid sequences. AAS64197-AAS94564 represent novel human
CC   diagnostic coding sequences of the invention.
CC   Note: The sequence data for this patent did not appear in the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 1082 BP; 193 A; 334 C; 353 G; 202 T; 0 other;

  Query Match          67.1%;  Score 923.6;  DB 23;  Length 1082;
  Best Local Similarity  96.5%;  Pred. No. 4.4e-161;
  Matches 1042;  Conservative    0;  Mismatches   24;  Indels   14;  Gaps    9;

Qy        299 CCCAGCCTCTGCCCGGGGCACCCCGGCCTTCCAGGCACGCCGGGCCACCATGGCAGCCAG 358
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CCCAGCCTCTGCCCGGGGCACCCCGGCCTTCCAGGCACGCCGGGCCACCATGGCAGCCAG 60

Qy        359 GGCTTGCCGGGCCGCGATGGCCGCGACGGCCGCGACGGCGCGCCCGGGGCTCCGGGAGAG 418
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GGCTTGCCGGGCCGCGATGGCCGCGACGGCCGCGACGGCGCGCCCGGGGCTCCGGGAGAG 120

Qy        419 AAAGGCGAGGGCGGGAGGCCGGGACTGCCGGGACCTCGAGGGGACCCCGGGCCGCGAGGA 478
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AAAGGCGAGGGCGGGAGGCCGGGACTGCCGGGACCTCGAGGGGACCCCGGGCCGCGAGGA 180

Qy        479 GAGGCGGGACCCGCGGGGCCCACCGGGCCTGCCGGGGAGTGCTCGGTGCCTCCGCGATCC 538
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GAGGCGGGACCCGCGGGGCCCACCGGGCCTGCCGGGGAGTGCTCGGTGCCTCCGCGATCC 240

Qy        539 GCCTTCAGCGCCAAGCGCTCCGAGAGCCGGGTGCCTCCGCCGTCTGACGCACCCTTGCCC 598
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GCCTTCAGCGCCAAGCGCTCCGAGAGCCGGGTGCCTCCGCCGTCTGACGCACCCTTGCCC 300

Qy        599 TTCGACCGCGTGCTGGTGAACGAGCAGGGACATTACGACGCCGTCACCGGCAAGTTCACC 658
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TTCGACCGCGTGCTGGTGAACGAGCAGGGACATTACGACGCCGTCACCGGCAAGTTCACC 360

Qy        659 TGCCAGGTGCCTGGGGTCTACTACTTCGCCGTCCATGCCACCGTCTACCGGGCCAGCCTG 718
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TGCCAGGTGCCTGGGGTCTACTACTTCGCCGTCCATGCCACCGTCTACCGGGCCAGCCTG 420

Qy        719 CAGTTTGATCTGGTGAAGAATGGCGAATCCATTGCCTCTTTCTTCCAGTTTTTCGGGGGG 778
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CAGTTTGATCTGGTGAAGAATGGCGAATCCATTGCCTCTTTCTTCCAGTTTTTCGGGGGG 480

Qy        779 TGGCCCAAGCCAGCCTCGCTCTCGGGGGGGGCCATGGTGAGGCTGGAGCCTGAGGACCAA 838
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TGGCCCAAGCCAGCCTCGCTCTCGGGGGGGGCCATGGTGAGGCTGGAGCCTGAGGACCAA 540

Qy        839 GTGTGGGTGCAGGTGGGTGTGGGTGACTACATTGGCATCTATGCCAGCATCAAGACAGAC 898
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GTGTGGGTGCAGGTGGGTGTGGGTGACTACATTGGCATCTATGCCAGCATCAAGACAGAC 600

Qy        899 AGCACCTTCTCCGGATTTCTGGTGTACTCCGACTGG---CACAGCTCCCCAGTCTTTGCT 955
              |||||||||||||||||||||||||||||||||||||  ||   | |||     || |||
Db        601 AGCACCTTCTCCGGATTTCTGGTGTACTCCGACTGGCATCAGCTCCCCCAGTCTTTAGCT 660

Qy        956 TAGTGCCC--ACTGCAAAGTGAGCTCATGCTCTC-ACTCCTAGAAGGAGGGTGTGAGGCT 1012
              ||||||||  ||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        661 TAGTGCCCCACTGCAAAAGTGAGCTCATGCTCTCAACTCCTAGAAGGAGGGTGTGAGGCT 720

Qy       1013 GACAACCAGGTCATCC-AGGAGGGCTGGCCCCCCTGGAATATTGTGAATGACTAGGGAGG 1071
              |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        721 GACAACCAGGTCATCCAAGGAGGGCTGGCCCCCCTGGAATATTGTGAATGACTAGGGAGG 780

Qy       1072 T--GGGGTAGAGCACTCTCC-GTCCTGCTGCTGGCAAGGAAT-GGGAACAGTGGCTGTCT 1127
              |  |||||||||||||||| |||||||||||||||||||||| ||||||||||||||||||
Db        781 TTGGGGGTAGAGCACTCTCCAGTCCTGCTGCTGGCAAGGAATGGGGAACAGTGGCTGTCT 840

Qy       1128 GCGATCAGGTCTGGCAGCATGGGGCAGTGGCTGGATTTCTGCCCAAGACCAGAGGAGTGT 1187
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GCGATCAGGTCTGGCAGCATGGGGCAGTGGCTGGATTTCTGCCCAAGACCAGAGGAGTGT 900

Qy       1188 GCTGTGCTGGCAAG-TGTAAGTCCCCCAGTTGCTCTGGTCCAGGAGCCCACGGTGGGGTG 1246
              |||||| ||| ||| |||||||||||||||||||| ||||||||||||||||||||||||
Db        901 GCTGTGGTGGGAAGTTGTAAGTCCCCCAGTTGCTTTGGTCCAGGAGCCCACGGTGGGGTG 960

Qy       1247 CTCTCTTCCTGGTCCTCTGCTTCTCTGGATCCTCCCCACCCCCTCCTGCTCCTGGGGCCG 1306
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 CTCTCTTCCTGGTCCTCTGCTTCTCTGGATCCTCCCCACCCCCTCCTGCTCCTGGGGCCG 1020

Qy       1307 GCCCTTTTCTCA--GAGATCACTCAATAAACCTAAGAACCCTCATAAAAAAAAAAAAAAA 1364
              ||||||||||||  |||||||||||||||||||||||||| ||||||||||||||||
Db       1021 GCCCTTTTCTCAAGAGATCCACTCAATAAACCTAAGAACCCTCAAAAAAAAAAAAAAAA 1080

Search completed: June 22, 2003, 03:43:14.
Job time : 355 secs
```

GenCore version 5.1.6


OM nucleic - nucleic search, using sw model

Run on: June 22, 2003, 03:22:25 ; Search time 1901 Seconds
(without alignments)
11731.292 Million cell updates/sec

Title: US-09-943-851A-41
Perfect score: 1377
Sequence: 1 gactagttctcttggagtct.....aaaaaaaaaaaaaaaaaaa 1377

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| | Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|---|
| | 1 | 767.8 | 55.8 | 869 | 13 | BI763193 | BI763193 603049929 |
| | 2 | 750.6 | 54.5 | 1154 | 13 | BM547549 | BM547549 AGENCOURT |
| | 3 | 732.8 | 53.2 | 1093 | 14 | BM924569 | BM924569 AGENCOURT |
| | 4 | 730.2 | 53.0 | 846 | 13 | BI458455 | BI458455 603198837 |
| c | 5 | 712.6 | 51.8 | 801 | 13 | BI490880 | BI490880 603031867 |
| | 6 | 672.2 | 48.8 | 724 | 13 | BI771879 | BI771879 603055280 |
| | 7 | 671 | 48.7 | 1045 | 14 | BM920874 | BM920874 AGENCOURT |
| | 8 | 660.4 | 48.0 | 792 | 13 | BI770921 | BI770921 603059737 |
| | 9 | 641.8 | 46.6 | 887 | 13 | BI490062 | BI490062 603031867 |
| | 10 | 631 | 45.8 | 1623 | 14 | BM926477 | BM926477 AGENCOURT |
| | 11 | 607.2 | 44.1 | 630 | 12 | BG706609 | BG706609 602672994 |
| c | 12 | 596 | 43.3 | 596 | 14 | BM893587 | BM893587 ij26g10.x |
| | 13 | 586.4 | 42.6 | 596 | 13 | BM544255 | BM544255 AGENCOURT |
| | 14 | 583.2 | 42.4 | 941 | 13 | BI821899 | BI821899 603035796 |
| c | 15 | 578 | 42.0 | 578 | 14 | BM893691 | BM893691 ij28e08.x |
| | 16 | 577.2 | 41.9 | 879 | 13 | BI820945 | BI820945 603033989 |
| c | 17 | 572.4 | 41.6 | 574 | 13 | BM021127 | BM021127 ie76e01.x |
| | 18 | 554 | 40.2 | 870 | 9 | AF451167 | AF451167 AF451167 |
| c | 19 | 551.4 | 40.0 | 553 | 12 | BF882978 | BF882978 IL3-ET011 |
| | 20 | 532 | 38.6 | 532 | 12 | BF724241 | BF724241 bx02e11.y |
| | 21 | 521 | 37.8 | 521 | 14 | BM966810 | BM966810 ij26g10.y |
| | 22 | 517 | 37.5 | 517 | 14 | BM966954 | BM966954 ij28e08.y |
| c | 23 | 507.8 | 36.9 | 574 | 10 | AW150070 | AW150070 xg48b05.x |
| | 24 | 507.6 | 36.9 | 844 | 10 | BE309370 | BE309370 601094718 |
| | 25 | 489.4 | 35.5 | 567 | 13 | BM021381 | BM021381 ie76e01.y |
| | 26 | 479 | 34.8 | 778 | 12 | BF098614 | BF098614 601750614 |
| | 27 | 463.4 | 33.7 | 867 | 13 | BI818537 | BI818537 603033057 |
| | 28 | 457.4 | 33.2 | 1349 | 14 | BQ673882 | BQ673882 AGENCOURT |
| c | 29 | 453.2 | 32.9 | 702 | 12 | BF055285 | BF055285 7j77g02.x |
| c | 30 | 452.4 | 32.9 | 454 | 9 | AI805087 | AI805087 tu33g05.x |
| c | 31 | 448.4 | 32.6 | 462 | 10 | AW070344 | AW070344 xa10d10.x |
| c | 32 | 440.2 | 32.0 | 445 | 14 | BM673416 | BM673416 UI-E-CR1- |
| | 33 | 438.4 | 31.8 | 842 | 13 | BI688017 | BI688017 603315786 |
| c | 34 | 426.4 | 31.0 | 449 | 13 | BM668611 | BM668611 UI-E-CK1- |
| | 35 | 424.4 | 30.8 | 442 | 14 | BM693336 | BM693336 UI-E-CK1- |
| c | 36 | 423.4 | 30.7 | 548 | 12 | BF906443 | BF906443 IL0-OT012 |
| | 37 | 421.6 | 30.6 | 630 | 12 | BF046067 | BF046067 BP250014A |
| c | 38 | 409.8 | 29.8 | 413 | 9 | AI083823 | AI083823 qf18c08.x |
| | 39 | 408.4 | 29.7 | 410 | 14 | BM707291 | BM707291 UI-E-CR1- |
| c | 40 | 405.4 | 29.4 | 408 | 9 | AI085548 | AI085548 qf24c09.x |
| c | 41 | 402.4 | 29.2 | 405 | 9 | AI083824 | AI083824 qf18c09.x |
| | 42 | 395.6 | 28.7 | 547 | 12 | BE722922 | BE722922 191347 MA |
| | 43 | 394.6 | 28.7 | 936 | 12 | BF540263 | BF540263 602052317 |
| c | 44 | 392 | 28.5 | 392 | 10 | BE138819 | BE138819 xw95h11.x |
| c | 45 | 387.2 | 28.1 | 392 | 10 | AW192716 | AW192716 xl50b08.x |

ALIGNMENTS

RESULT 1
BI763193
LOCUS BI763193 869 bp mRNA linear EST 25-SEP-2001
DEFINITION 603049929F1 NIH_MGC_116 Homo sapiens cDNA clone IMAGE:5190222 5', mRNA sequence.
ACCESSION BI763193
VERSION BI763193.1 GI:15754771
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 869)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM11475 row: i column: 07
High quality sequence stop: 777.
FEATURES Location/Qualifiers
source 1. .869

```
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5190222"
                     /clone_lib="NIH_MGC_116"
                     /lab_host="DH10B"
                     /note="Organ: pooled colon, kidney, stomach; Vector:
                     pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
                     source anonymous pool of 3 colons, age 26 yo male, 49 yo
                     female, 71 yo male colon; 46 yo male kidney, and pool of 2
                     stomachs, 62 yo male and 70 yo female.  Library is
                     oligo-dT primed and  directionally cloned (EcoRV site is
                     destroyed upon cloning).  Average insert size 1.4 kb,
                     insert size range 1-3 kb.  Library is normalized and
                     enriched for full-length clones and was constructed by C.
                     Gruber  (Invitrogen).  Research Genetics tracking code
                     023. Note: this is a NIH_MGC Library."
BASE COUNT       133 a    277 c    339 g    120 t
ORIGIN

  Query Match             55.8%;  Score 767.8;  DB 13;  Length 869;
  Best Local Similarity   97.7%;  Pred. No. 7.9e-96;
  Matches  853;  Conservative    0;  Mismatches   12;  Indels    8;  Gaps    7;

Qy         11 CTTGGAGTCTGGGAGGAGGAAAGCGGAGCCGGCAGGGAGCGAACCAGGACTGGGGTGACG 70
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CTTGGAGTCTGGGAGGAGGAAAGCGGAGCCGGCAGGGAGCGAACCAGGACTGGGGTGAGG 60

Qy         71 GCAGGGCAGGGGGCGCCTGGCCGGGGAGAAGCGCGGGGGCTGGAGCACCACCAACTGGAG 130
               |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GCAGGGCAGGTGGCGCCTGGCCGGGGAGAAGCGCGGGGGCTGGAGCACCACCAACTGGAG 120

Qy        131 GGTCCGGAGTAGCGAGCGCCCCGAAGGAGGCCATCGGGGAGCCGGGAGGGGGGACTGCGA 190
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GGTCCGGAGTAGCGAGCGCCCCGAAGGAGGCCATCGGGGAGCCGGGAGGGGGGACTGCGA 180

Qy        191 GAGGACCCCGGCGTCCGGGCTCCCGGTGCCAGCGCTATGAGGCCACTCCTCGTCCTGCTG 250
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GAGGACCCCGGCGTCCGGGCTCCCGGTGCCAGCGCTATGAGGCCACTCCTCGTCCTGCTG 240

Qy        251 CTCCTGGGCCTGGCGGCCGGCTCGCCCCCACTGGACGACAACAAGATCCCCAGCCTCTGC 310
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CTCCTGGGCCTGGCGGCCGGCTCGCCCCCACTGGACGACAACAAGATCCCCAGCCTCTGC 300

Qy        311 CCGGGGCACCCCGGCCTTCCAGGCACGCCGGGCCACCATGGCAGCCAGGGCTTGCCGGGC 370
               |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db        301 CCGGGGCACCCCGGCCTTCCAGGCACGCCGGGCCACCATGGCAGCCA-GGCTTGCCGGGC 359

Qy        371 CGCGATGGCCGCGACGGCCGCGACGGCGCGCCCGGGGCTCCGGGAGAGAAAGGCGAGGGC 430
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        360 CGCGATGGCCGCGACGGCCGCGACGGCGCGCCCGGGGCTCCGGGAGAGAAAGGCGAGGGC 419

Qy        431 GGGAGGCCGGGACTGCCGGGACCTCGAGGGGACCCCGGGCCGCGAGGAGAGGCGGGACCC 490
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        420 GGGAGGCCGGGACTGCCGGGACCTCGAGGGGACCCCGGGCCGCGAGGAGAGGCGGGACCC 479

Qy        491 GCGGGGCCCACCGGGCCTGCCGGGGAGTGCTCGGTGCCTCCGCGATCCGCCTTCAGCGCC 550
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        480 GCGGGGCCCACCGGGCCTGCCGGGGAGTGCTCGGTGCCTCCGCGATCCGCCTTCAGCGCC 539

Qy        551 AAGCGCTCCGAGAGCC-GGGTGCCTCCGCCGTCTGACGCACCCTTGCCCTTCGACCGCGT 609
               |||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||
Db        540 AAGCGCTCCGAGAGCCGGGGTGCCTCCGCCGTCTGACGCACCCTTGCCCTTCGACCGCGT 599

Qy        610 GCTGGTG-AACGAGCAGGGACATTACGACGCCG-TCACCGGCAAGTTCACCTGCCAGGTG 667
               ||||||| |||||||||||||||||||||||| |||||||||||||||||||||||||||
Db        600 GCTGGTGAAACGAGCAGGGACATTACGACGCCGTTCACCGGCAAGTTCACCTGCCAGGTG 659

Qy        668 CCTGGGGTCTACTACTTCGCCGTCCATGCCACCGTCTACCGGGCCAGCCTGCAGTTTGAT 727
               |||||||||||||||||||||||||||||||||||||||||||||||| |||||||||||
Db        660 CCTGGGGTCTACTACTTCGCCGTCCATGCCACCGTCTACCGGGCCAGCCTGCAGGTTGAT 719

Qy        728 CTGGTGAAGAATGGCGAATCCATTGCCTCTTTCTTCCAGTTTTTCGGGGGGTGGCCCAAG 787
               |||||||||||||||||||||||| ||| |||||||||||| |||||||||||  |
Db        720 CTGGTGAAGAATGGCGAATCCATTGGCTC-TTCTTCCAGTTTTACGGGGGGTGGCC--AA 776

Qy        788 CCAGCCTCGCTCTCGGGGGGGGCCATGGTGAGGCTGGAGCCTGAGGACCAAGTGTGGGTG 847
               |   |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        777 GCCAGCTCGCTCTCGGGGGGGGCCATGGTGAGGCTGGAGCCTGAGGACCAAGTGTGGGTG 836

Qy        848 CAGGTGGGTG-TGGGTGACTACATTGGCATCTA 879
               ||||| |||| |||||||||||| ||||||||||
Db        837 CAGGTTGGTGTTGGGTGACTACGTTGGCATCTA 869


RESULT 2
BM547549
LOCUS       BM547549                1154 bp    mRNA    linear   EST 20-FEB-2002
DEFINITION  AGENCOURT_6507694 NIH_MGC_125 Homo sapiens cDNA clone IMAGE:5724522
            5', mRNA sequence.
ACCESSION   BM547549
VERSION     BM547549.1  GI:18781410
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1154)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Invitrogen
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM12713  row: o  column: 19
            High quality sequence stop: 591.
FEATURES             Location/Qualifiers
     source          1. .1154
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5724522"
                     /clone_lib="NIH_MGC_125"
                     /lab_host="DH10B"
                     /note="Organ: ovary (pool of 3); Vector: pCMV-SPORT6;
                     Site_1: EcoRV (destroyed); Site_2: NotI; RNA source pool
                     of three ovaries, from females ranging in age from 38 to
                     49 yo. Library is oligo-dT primed and directionally cloned
                     (EcoRV site is destroyed upon cloning). Average insert
                     size 2.1 kb, insert size range 1-3.5 kb. Library is
                     normalized and enriched for full-length clones and was
                     constructed by C. Gruber (Invitrogen). Research Genetics
                     tracking code 036."
BASE COUNT       178 a    382 c    401 g    193 t
ORIGIN

  Query Match             54.5%;  Score 750.6;  DB 13;  Length 1154;
  Best Local Similarity   93.3%;  Pred. No. 1.5e-93;
  Matches  818;  Conservative    0;  Mismatches   54;  Indels    5;  Gaps    3;

Qy         40 CGGCAGGGAGCGAACCAGGACTGGGGTGACGGCAGGGCAGGGGGCGCCTGGCCGGGGAGA 99
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CGGCAGGGAGCGAACCAGGACTGGGGTGACGGCAGGGCAGGGGGCGCCTGGCCGGGGAGA 60

Qy        100 AGCGCGGGGGCTGGAGCACCACCAACTGGAGGGTCCGGAGTAGCGAGCGCCCCGAAGGAG 159
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AGCGCGGGGGCTGGAGCACCACCAACTGGAGGGTCCGGAGTAGCGAGCGCCCCGAAGGAG 120

Qy        160 GCCATCGGGGAGCCGGGAGGGGGGACTGCGAGAGGACCCCGGCGTCCGGGCTCCCGGTGC 219
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GCCATCGGGGAGCCGGGAGGGGGGACTGCGAGAGGACCCCGGCGTCCGGGCTCCCGGTGC 180
```

```
Qy        220 CAGCGCTATGAGGCCACTCCTCGTCCTGCTGCTCCTGGGCCTGGCGGCCGGCTCGCCCCC 279
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CAGCGCTATGAGGCCACTCCTCGTCCTGCTGCTCCTGGGCCTGGCGGCCGGCTCGCCCCC 240

Qy        280 ACTGGACGACAACAAGATCCCCAGCCTCTGCCCGGGGCACCCCGGCCTTCCAGGCACGCC 339
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ACTGGACGACAACAAGATCCCCAGCCTCTGCCCGGGGCACCCCGGCCTTCCAGGCACGCC 300

Qy        340 GGGCCACCATGGCAGCCAGGGCTTGCCGGGCCGCGATGGCCGCGACGGCCGCGACGGCGC 399
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GGGCCACCATGGCAGCCAGGGCTTGCCGGGCCGCGATGGCCGCGACGGCCGCGACGGCGC 360

Qy        400 GCCCGGGGCTCCGGGAGAGAAAGGCGAGGGCGGGAGGCCGGGACTGCCGGGACCTCGAGG 459
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GCCCGGGGCTCCGGGAGAGAAAGGCGAGGGCGGGAGGCCGGGACTGCCGGGACCTCGAGG 420

Qy        460 GGACCCCGGGCCGCGAGGAGAGGCGGGACCCGCGGGGCCCACCGGGCCTGCCGGGGAGTG 519
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GGACCCCGGGCCGCGAGGAGAGGCGGGACCCGCGGGGCCCACCGGGCCTGCCGGGGAGTG 480

Qy        520 CTCGGTGCCTCCGCGATCCGCCTTCAGCGCCAAGCGCTCCGAGAGCCGGGTGCCTCCGCC 579
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CTCGGTGCCTCCGCGATCCGCCTTCAGCGCCAAGCGCTCCGAGAGCCGGGTGCCTCCGCC 540

Qy        580 GTCTGACGCACCCTTGCCCTTCGACCGCGTGCTGGTGAACGAGCAGGGACATTACGACGC 639
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GTCTGACGCACCCTTGCCCTTCGACCGCGTGCTGGTGAACGAGCAGGGACATTACGACGC 600

Qy        640 CGTCACCGGCAAGTTCACCTGCCAGGTGCCTGGGGTCTACTACTTCGCCGTCCATGCCAC 699
              ||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db        601 CGTCACCGGCAAGTTCACCTGCCAGGTGCCTGGGGTCTACTACTTCGGCGTCCATGCCAC 660

Qy        700 CGTCTACCGGGCCAGCCTGCAGTTTGATCTGGTGAAGAATGGCGAATCCATTGCCTCTTT 759
              ||||||||||||||||||||||||||| || |||||||||||||||| ||||||||||
Db        661 CGTCTACCGGGCCAGCCTGCAGTTTGATCCGGGGAAGAATGGCGAATCC-TTGCCTCTTT 719

Qy        760 CTTCCAGTTTTTCGGGGGGTGGCCCAAGCCAGCCTCGCTCTCGGGGGGGGCCATGGTGAG 819
              ||||  ||||  |||||| | || |  ||  |||    ||||||||||||  ||||||
Db        720 CTTCAGTTTTTCGGGGGGGGGCCCAAGCCCGCCCTGCTCCTCGGGGGGGGCACTGGTGAA 779

Qy        820 GCTGG-AGCCTGAGGACCAAGTGTGGGTGCAGGTGGGTGTGGG---TGACTACATTGGCA 875
              | ||| | ||||||||||||||||||| |||   ||||  | |      || | | || |
Db        780 GGTGGAACCCTGAGGACCAAGTGTGGGGGCACCGGGGTCGGAGGTTGACCTCCCTGGGGA 839

Qy        876 TCTATGCCAGCATCAAGACAGACAGCACCTTCTCCGG 912
              ||||||||||     ||  || |    | | ||| |
Db        840 CCTATGCCAGCGGTCCGAACGAAACAGACCTTTCCTG 876
```

```
RESULT 3
BM924569
LOCUS       BM924569                1093 bp    mRNA    linear   EST 12-MAR-2002
DEFINITION  AGENCOURT_6767665 NIH_MGC_116 Homo sapiens cDNA clone IMAGE:5760942
            5', mRNA sequence.
ACCESSION   BM924569
VERSION     BM924569.1  GI:19374948
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1093)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Life Technologies, Inc.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM12808  row: m  column: 07
            High quality sequence stop: 590.
FEATURES             Location/Qualifiers
     source          1. .1093
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5760942"
                     /clone_lib="NIH_MGC_116"
                     /lab_host="DH10B"
                     /note="Organ: pooled colon, kidney, stomach; Vector:
                     pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
                     source anonymous pool of 3 colons, age 26 yo male, 49 yo
                     female, 71 yo male colon; 46 yo male kidney, and pool of 2
                     stomachs, 62 yo male and 70 yo female.  Library is
                     oligo-dT primed and  directionally cloned (EcoRV site is
                     destroyed upon cloning).  Average insert size 1.4 kb,
                     insert size range 1-3 kb.  Library is normalized and
                     enriched for full-length clones and was constructed by C.
                     Gruber  (Invitrogen).  Research Genetics tracking code
                     023. Note: this is a NIH_MGC Library."
BASE COUNT      197 a    339 c    356 g    198 t      3 others
ORIGIN
```

Query Match 53.2%; Score 732.8; DB 14; Length 1093;
Best Local Similarity 96.8%; Pred. No. 4.1e-91;
Matches 757; Conservative 0; Mismatches 24; Indels 1; Gaps 1;

```
Qy        331 AGGCACGCCGGGCCACCATGGCAGCCAGGGCTTGCCGGGCCGCGATGGCCGCGACGGCCG 390
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AGGCACGCCGGGCCACCATGGCAGCCAGGGCTTGCCGGGCCGCGATGGCCGCGACGGCCG 60

Qy        391 CGACGGCGCGCCCGGGGCTCCGGGAGAGAAAGGCGAGGGCGGGAGGCCGGGACTGCCGGG 450
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CGACGGCGCGCCCGGGGCTCCGGGAGAGAAAGGCGAGGGCGGGAGGCCGGGACTGCCGGG 120

Qy        451 ACCTCGAGGGGACCCCGGGCCGCGAGGAGAGGCGGGACCCGCGGGGCCCACCGGGCCTGC 510
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACCTCGAGGGGACCCCGGGCCGCGAGGAGAGGCGGGACCCGCGGGGCCCACCGGGCCTGC 180

Qy        511 CGGGGAGTGCTCGGTGCCTCCGCGATCCGCCTTCAGCGCCAAGCGCTCCGAGAGCCGGGT 570
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CGGGGAGTGCTCGGTGCCTCCGCGATCCGCCTTCAGCGCCAAGCGCTCCGAGAGCCGGGT 240

Qy        571 GCCTCCGCCGTCTGACGCACCCTTGCCCTTCGACCGCGTGCTGGTGAACGAGCAGGGACA 630
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GCCTCCGCCGTCTGACGCACCCTTGCCCTTCGACCGCGTGCTGGTGAACGAGCAGGGACA 300

Qy        631 TTACGACGCCGTCACCGGCAAGTTCACCTGCCAGGTGCCTGGGGTCTACTACTTCGCCGT 690
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TTACGACGCCGTCACCGGCAAGTTCACCTGCCAGGTGCCTGGGGTCTACTACTTCGCCGT 360

Qy        691 CCATGCCACCGTCTACCGGGCCAGCCTGCAGTTTGATCTGGTGAAGAATGGCGAATCCAT 750
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CCATGCCACCGTCTACCGGGCCAGCCTGCAGTTTGATCTGGTGAAGAATGGCGAATCCAT 420

Qy        751 TGCCTCTTTCTTCCAGTTTTTCGGGGGGTGGCCCAAGCCAGCCTCGCTCTCGGGGGGGGC 810
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TGCCTCTTTCTTCCAGTTTTTCGGGGGGTGGCCCAAGCCAGCCTCGCTCTCGGGGGGGGC 480

Qy        811 CATGGTGAGGCTGGAGCCTGAGGACCAAGTGTGGGTGCAGGTGGGTGTGGGTGACTACAT 870
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CATGGTGAGGCTGGAGCCTGAGGACCAAGTGTGGGTGCAGGTGGGTGTGGGTGACTACAT 540

Qy        871 TGGCATCTATGCCAGCATCAAGACAGACAGCACCTTCTCCGGATTTCTGGTGTACTCCGA 930
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TGGCATCTATGCCAGCATCAAGACAGACAGCACCTTCTCCGGATTTCTGGTGTACTCCGA 600

Qy        931 CTGGCACAGCTCCCCAGTCTTTGCTTAGTGCCCACTGCAAAGTGAGCTCATGCTCTCACT 990
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CTGGCACAGCTCCCCAGTCTTTGCTTAGTGCCCACTGCAAAGTGAGCTCATGCTCTCACT 660

Qy        991 CCTAG-AAGGAGGGTGTGAGGCTGACAACCAGGTCATCCAGGAGGGCTGGCCCCCCTGGA 1049
```

```
Db        661 CCTAGAAAGGAGGGTGTGAGGCTGACAACCAGGTCATCCAGGAGGGCTGGGCCCCCTGGA 720

Qy       1050 ATATTGTGAATGACTAGGGAGGTGGGGTAGAGCACTCTCCGTCCTGCTGCTGGCAAGGAA 1109

Db        721 ATATTGTGAATGACTAGGGAAGGGGGGGTANAANCACTCTCCGTCCCTGCTGCTTGGCAA 780

Qy       1110 TG 1111

Db        781 GG 782
```

RESULT 4
BI458455

| | |
|---|---|
| LOCUS | BI458455 846 bp mRNA linear EST 21-AUG-2001 |
| DEFINITION | 603198837F1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:5278184 5', mRNA sequence. |
| ACCESSION | BI458455 |
| VERSION | BI458455.1 GI:15249111 |
| KEYWORDS | EST. |
| SOURCE | human. |
| ORGANISM | Homo sapiens<br>Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. |
| REFERENCE | 1. (bases 1 to 846) |
| AUTHORS | NIH-MGC http://mgc.nci.nih.gov/. |
| TITLE | National Institutes of Health, Mammalian Gene Collection (MGC) |
| JOURNAL | Unpublished (1999) |
| COMMENT | Contact: Robert Strausberg, Ph.D.<br>Email: cgapbs-r@mail.nih.gov<br>Tissue Procurement: Miklos Palkovits, M.D., Ph.D.<br>cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki Toshiyuki and Piero Carninci (RIKEN)<br>cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)<br>DNA Sequencing by: Incyte Genomics, Inc.<br>Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov<br>Plate: LLAM11702 row: j column: 09<br>High quality sequence stop: 815. |

FEATURES Location/Qualifiers

source 1..846
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5278184"
/clone_lib="NIH_MGC_96"
/tissue_type="hypothalamus"
/lab_host="DH10B"
/note="Organ: brain; Vector: pBluescriptR (modified pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag ); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTTVN-3', size-selected for average insert size 2.3 kb and normalized to ROT 5. This is a primary library enriched for full-length clones and constructed using the Cap-trapper method (Carninci, in preparation). Library constructed by M. Brownstein (NIMH/NHGRI, National Institutes of Health). Note: this is a NIH_MGC Library."

BASE COUNT 129 a 271 c 330 g 115 t 1 others
ORIGIN

Query Match 53.0%; Score 730.2; DB 13; Length 846;
Best Local Similarity 96.1%; Pred. No. 1.1e-90;
Matches 803; Conservative 0; Mismatches 24; Indels 9; Gaps 5;

```
Qy          7 TTCTCTTGGAGTCTGGGAGGAGGAAAGCGGAGCCGGCAGGGAGCGAACCAGGACTGGGGT 66

Db         19 TCCTCTTGGAGTCTGGGAGGAGGAAAGCGGAGCCGGCAGGGAGCGAACCAGGACTGGGGT 78

Qy         67 GACGGCAGGGCAGGGGGCGCCTGGCCGGGGAGAAGCGCGGGGGCTGGAGCACCACCAACT 126

Db         79 GACGGCAGGGCAGGGGGCGCCTGGCCGGGGAGAAGCGCGGGGGCTGGAGCACCACCAACT 138

Qy        127 GGAGGGTCCGGAGTAGCGAGCGCCCCGAAGGAGGCCATCGGGGAGCCGGGAGGGGGGACT 186

Db        139 GGAGGGTCCGGAGTAGCGAGCGCCCCGAAGGAGGCCATCGGGGAGCCGGGAGGGGGGACT 198

Qy        187 GCGAGAGGACCCCGGCGTCCGGGCTCCCGGTGCCAGCGCTATGAGGCCACTCCTCGTCCT 246

Db        199 GCGAGAGGACCCCGGCGTCCGGGCTCCCGGTGCCAGCGCTATGAGGCCACTCCTCGTCCT 258

Qy        247 GCTGCTCCTGGGCCTGGCGGCCGGCTCGCCCCCACTGGACGACAACAAGATCCCCAGCCT 306

Db        259 GCTGCTCCTGGGCCTGGCGGCCGGCTCGCCCCCACTGGACGACAACAAGATCCCCAGCCT 318

Qy        307 CTGCCCGGGGCACCCCGGCCTTCCAGGCACGCCGGGCCACCATGGCAGCCAGGGCTTGCC 366

Db        319 CTGCCCGGGGCACCCCGGCCTTCCAGGCACGCCGGGCCACCATGGCAGCCGGGGCTTGCC 378

Qy        367 GGGCCGCGATGGCCGCGACGGCCGCGACGGCGCGCCCGGGGCTCCGGGAGAGAAAGGCGA 426

Db        379 GGGCCGCGATGGCCGCGACGGCCGCGACGGCGCGCCCGGGGCTCCGGGAGAGAAAGGCGA 438

Qy        427 GGGCGGGAGGCCGGGACTGCCGGGACCTCGAGGGGACCCCGGGCCGCGAGGAGAGGCGGG 486

Db        439 GGGCGGGAGGCCGGGACTGCCGGGACCTCGAGGGGACCCCGGGCCGCGAGGAGAGGCGGG 498

Qy        487 ACCCGCGGGGCCCACCGGGCCTGCCGGGGAGTGCTCGGTGCCTCCGCGATCCGCCTTCAG 546

Db        499 ACCCGCGGGGCCCA-CGGGCCTGCCGGGGAGTGCTCGGTGCCTCCGCGATCCGCCTTCAG 557

Qy        547 CGCCAAGCGCTCCGAGAGCCGGGTGCCTCCGCCGTCTGACGCACCCTTGCCCTTCGACCG 606

Db        558 CGCCAAGCGCTCCGAGAGCCGGGTGCCTCCGCCGTCTGACGCACCCTTGCCCTTCGACCG 617

Qy        607 CGTGCTGGTGAACGAGCAGGGACATTACGACGCCGTCACCGGCAAGTTCACCT-GCCAGG 665

Db        618 CGTGCTGGTGAACGAGCAGGGACATTACGACGCCGTCACCGGCAAGTTCACCTGGCCAGG 677

Qy        666 TGCCTGGGGTCTACTACTTCGCCGTCCATGCCACCGTCTACCGGGCCAGCCTGCAGTTTG 725

Db        678 TGCCTGGGGTCTACTACTTCGNCGTCCATGCCACCGTCTACCGGGCCAG-CTGCAGTTTG 736

Qy        726 ATCTGGTGAAGAATGGCGAATCCATTGCCTCTTTCTTCCAGTTTTTCGGGGGGTGGCCCA 785

Db        737 ATCTGGTGAAGAATGG-GAATCCATTGCCTCTTTCTTCCAGTTTTCGGGGGGGTGGC--- 792

Qy        786 AGCCAGCCTCGCTCTCGGGGGGGGCCATGGTGAGGCTGGAGCCTGAGGACCAAGTG 841

Db        793 --CAGGCAAGCTCGTCTCGGGGGGGCATGTTGAGGCTGGAGCTTGAGACCAGTGTG 846
```

RESULT 5
BI490880/c

| | |
|---|---|
| LOCUS | BI490880 801 bp mRNA linear EST 28-AUG-2001 |
| DEFINITION | 603031867T1 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5172880 3', mRNA sequence. |
| ACCESSION | BI490880 |
| VERSION | BI490880.1 GI:15330108 |
| KEYWORDS | EST. |
| SOURCE | human. |
| ORGANISM | Homo sapiens<br>Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. |
| REFERENCE | 1 (bases 1 to 801) |
| AUTHORS | NIH-MGC http://mgc.nci.nih.gov/. |
| TITLE | National Institutes of Health, Mammalian Gene Collection (MGC) |
| JOURNAL | Unpublished (1999) |
| COMMENT | Contact: Robert Strausberg, Ph.D.<br>Email: cgapbs-r@mail.nih.gov<br>Tissue Procurement: Life Technologies, Inc.<br>cDNA Library Preparation: Life Technologies, Inc.<br>cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)<br>DNA Sequencing by: Incyte Genomics, Inc.<br>Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov |

```
            Plate: LLAM11430  row: f  column: 17
            High quality sequence start: 8
            High quality sequence stop: 738.
FEATURES              Location/Qualifiers
     source           1. .801
                      /organism="Homo sapiens"
                      /db_xref="taxon:9606"
                      /clone="IMAGE:5172880"
                      /clone_lib="NIH_MGC_115"
                      /lab_host="DH10B"
                      /note="Organ: pooled brain, lung, testis; Vector:
                      pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
                      source anonymous pool of 6 male brains, age range 23-27; 1
                      male lung, age 27; and 1 male testis, age 69.  Library is
                      oligo-dT primed and directionally cloned (EcoRV site is
                      destroyed upon  cloning).  Average insert size 1.8 kb,
                      insert size range 1-3 kb. Library is normalized and
                      enriched for full-length clones and was  constructed by C.
                      Gruber (Invitrogen).  Research Genetics tracking  code
                      021. Note: this is a NIH_MGC Library."
BASE COUNT       180 a    251 c    237 g    133 t
ORIGIN

  Query Match             51.8%;  Score 712.6;  DB 13;  Length 801;
  Best Local Similarity   98.2%;  Pred. No. 2.7e-88;
  Matches  773;  Conservative    0;  Mismatches    9;  Indels    5;  Gaps    5;

Qy        527 CCTCCGCGATCCGCCTTCAGC-GCCAAGCGCTCCGAGAGCCGGGTGCCTCCGCCGTCTGA 585
              ||| |||||||||||||||| ||||||||||||||||||||| |||||||||||||||
Db        786 CCTTCGCGATCCGCCTTCAGCGGCCAAGCGCTCCGAGAGCCGGGTTCCTCCGCCGTCTGA 727

Qy        586 CGCACCCTTGCCCTTCGACCGCGTGCTGGTGAACGAG-CAGGGACATTACGACGCCGTCA 644
              |||||| ||| |||||||||||||||||||||||||| ||||||||||||||||||||||
Db        726 CGCACCTTTG-CCTTCGACCGCGTGCTGGTGAACGAGCCAGGGACATTACGACGCCGTCA 668

Qy        645 CCGGC-AAGTTCACCTGCCAGGTGCCTGGGGTCTACTACTTCGCCGTCCATGCCACCGTC 703
              ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        667 CCGGCAAAGTTCACCTGCCAGGTGCCTGGGGTCTACTACTTCGCCGTCCATGCCACCGTC 608

Qy        704 TACCGGGCCAGCCTGCAGTTTGATCT-GGTGAAGAATGGCGAATCCATTGCCTCTTTCTT 762
              |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db        607 TACCGGGCCAGCCTGCAGTTTGATCTCGGTGAAGAATGGCGAATCCATTGCCTCTTTCTT 548

Qy        763 CCAGTTTTTCGGGGGGTGGCCCAAGCCAGCCTCGCTCTCGGGGGGGGCCATGGTGAGGCT 822
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        547 CCAGTTTTTCGGGGGGTGGCCCAAGCCAGCCTCGCTCTCGGGGGGGGCCATGGTGAGGCT 488

Qy        823 GGAGCCTGAGGACCAAGTGTGGGTGCAGGTGGGTGTGGGTGACTACATTGGCATCTATGC 882
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        487 GGAGCCTGAGGACCAAGTGTGGGTGCAGGTGGGTGTGGGTGACTACATTGGCATCTATGC 428

Qy        883 CAGCATCAAGACAGACAGCACCTTCTCCGGATTTCTGGTGTACTCCGACTGGCACAGCTC 942
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        427 CAGCATCAAGACAGACAGCACCTTCTCCGGATTTCTGGTGTACTCCGACTGGCACAGCTC 368

Qy        943 CCCAGTCTTTGCTTAGTGCCCACTGCAAAGTGAGCTCATGCTCTCACTCCTAGAAGGAGG 1002
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        367 CCCAGTCTTTGCTTAGTGCCCACTGCAAAGTGAGCTCATGCTCTCACTCCTAGAAGGAGG 308

Qy       1003 GTGTGAGGCTGACAACCAGGTCATCCAGGAGGGCTGGCCCCCCTGGAATATTGTGAATGA 1062
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        307 GTGTGAGGCTGACAACCAGGTCATCCAGGAGGGCTGGCCCCCCTGGAATATTGTGAATGA 248

Qy       1063 CTAGGGAGGTGGGGTAGAGCACTCTCCGTCCTGCTGCTGGCAAGGAATGGGAACAGTGGC 1122
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        247 CTAGGGAGGTGGGGTAGAGCACTCTCCGTCCTGCTGCTGGCAAGGAATGGGAACAGTGGC 188

Qy       1123 TGTCTGCGATCAGGTCTGGCAGCATGGGGCAGTGGCTGGATTTCTGCCCAAGACCAGAGG 1182
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        187 TGTCTGCGATCAGGTCTGGCAGCATGGGGCAGTGGCTGGATTTCTGCCCAAGACCAGAGG 128

Qy       1183 AGTGTGCTGTGCTGGCAAGTGTAAGTCCCCCAGTTGCTCTGGTCCAGGAGCCCACGGTGG 1242
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        127 AGTGTGCTGTGCTGGCAAGTGTAAGTCCCCCAGTTGCTCTGGTCCAGGAGCCCACGGTGG 68

Qy       1243 GGTGCTCTCTTCCTGGTCCTCTGCTTCTCTGGATCCTCCCCACCCCCTCCTGCTCCTGGG 1302
              ||||||||||||||||||||||||||||||||||||||||  ||||||||||||||||||
Db         67 GGTGCTCTCTTCCTGGTCCTCTGCTTCTCTGGATCCTCATAACCCCCTCCTGCTCCTGGG 8

Qy       1303 GCCGGCC 1309
              |  | ||
Db           7 GGGGCCC 1


RESULT 6
BI771879
LOCUS       BI771879                724 bp    mRNA    linear   EST 25-SEP-2001
DEFINITION  603055280F1 NIH_MGC_122 Homo sapiens cDNA clone IMAGE:5204971 5',
            mRNA sequence.
ACCESSION   BI771879
VERSION     BI771879.1  GI:15763457
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 724)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Life Technologies, Inc.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM11513  row: o  column: 20
            High quality sequence stop: 672.
FEATURES              Location/Qualifiers
     source           1. .724
                      /organism="Homo sapiens"
                      /db_xref="taxon:9606"
                      /clone="IMAGE:5204971"
                      /clone_lib="NIH_MGC_122"
                      /lab_host="DH10B"
                      /note="Organ: pooled lung and spleen; Vector: pCMV-SPORT6;
                      Site_1: NotI; Site_2: EcoRV (destroyed); RNA source
                      anonymous pool of 24 week female lung, 16 week female
                      spleen, and 20-22 week male spleens.  Library is oligo-dT
                      primed and directionally cloned (EcoRV site is destroyed
                      upon  cloning).  Average insert size 1.4 kb, insert size
                      range 1-3 kb. Library is normalized and enriched for
                      full-length clones and was  constructed by C. Gruber
                      (Invitrogen).  Research Genetics tracking  code 026. Note:
                      this is a NIH_MGC Library."
BASE COUNT       176 a    184 c    211 g    152 t      1 others
ORIGIN

  Query Match             48.8%;  Score 672.2;  DB 13;  Length 724;
  Best Local Similarity   99.4%;  Pred. No. 9.2e-83;
  Matches  674;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy        700 CGTCTACCGGGCCAGCCTGCAGTTTGATCTGGTGAAGAATGGCGAATCCATTGCCTCTTT 759
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          2 CGTCTACCGGGCCAGCCTGCAGTTTGATCTGGTGAAGAATGGCGAATCCATTGCCTCTTT 61

Qy        760 CTTCCAGTTTTTCGGGGGGTGGCCCAAGCCAGCCTCGCTCTCGGGGGGGGCCATGGTGAG 819
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         62 CTTCCAGTTTTTCGGGGGGTGGCCCAAGCCAGCCTCGCTCTCGGGGGGGGCCATGGTGAG 121

Qy        820 GCTGGAGCCTGAGGACCAAGTGTGGGTGCAGGTGGGTGTGGGTGACTACATTGGCATCTA 879
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        122 GCTGGAGCCTGAGGACCAAGTGTGGGTGCAGGTGGGTGTGGGTGACTACATTGGCATCTA 181
```

```
Qy        880 TGCCAGCATCAAGACAGACAGCACCTTCTCCGGATTTCTGGTGTACTCCGACTGGCACAG 939
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        182 TGCCAGCATCAAGACAGACAGCACCTTCTCCGGATTTCTGGTGTACTCCGACTGGCACAG 241

Qy        940 CTCCCCAGTCTTTGCTTAGTGCCCACTGCAAAGTGAGCTCATGCTCTCACTCCTAGAAGG 999
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        242 CTCCCCAGTCTTTGCTTAGTGCCCACTGCAAAGTGAGCTCATGCTCTCACTCCTAGAAGG 301

Qy       1000 AGGGTGTGAGGCTGACAACCAGGTCATCCAGGAGGGCTGGCCCCCCCTGGAATATTGTGAA 1059
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        302 AGGGTGTGAGGCTGACAACCAGGTCATCCAGGAGGGCTGGCCCCCCTGGAATATTGTGAA 361

Qy       1060 TGACTAGGGAGGTGGGGTAGAGCACTCTCCGTCCTGCTGCTGGCAAGGAATGGGAACAGT 1119
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        362 TGACTAGGGAGGTGGGGTAGAGCACTCTCCGTCCTGCTGCTGGCAAGGAATGGGAACAGT 421

Qy       1120 GGCTGTCTGCGATCAGGTCTGGCAGCATGGGGCAGTGGCTGGATTTCTGCCCAAGACCAG 1179
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        422 GGCTGTCTGCGATCAGGTCTGGCAGCATGGGGCAGTGGCTGGATTTCTGCCCAAGACCAG 481

Qy       1180 AGGAGTGTGCTGTGCTGGCAAGTGTAAGTCCCCCAGTTGCTCTGGTCCAGGAGCCCACGG 1239
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        482 AGGAGTGTGCTGTGCTGGCAAGTGTAAGTCCCCCAGTTGCTCTGGTCCAGGAGCCCACGG 541

Qy       1240 TGGGGTGCTCTCTTCCTGGTCCTCTGCTTCTCTGGATCCTCCCCACCCCCTCCTGCTCCT 1299
              ||||||||||||||||||||||||||||||||||||||||||  ||||||||||||||||
Db        542 TGGGGTGCTCTCTTCCTGGTCCTCTGCTTCTCTGGATCCTCCCCNACCCCTCCTGCTCCT 601

Qy       1300 GGGGCCGGCCCTTTTCTCAGAGATCACTCAATAAACCTAAGAACCCTCATAAAAAAAAAA 1359
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        602 GGGGCCGGCCCTTTTCTCAGAGATCACTCAATAAACCTAAGAACCCTCAAAAAAAAAAAA 661

Qy       1360 AAAAAAAAAAAAAAAAA 1377
              ||||||||||||||| |||
Db        662 AAAAAAAAAAAAAAGAAA 679
```

RESULT 7
BM920874
LOCUS BM920874 1045 bp mRNA linear EST 12-MAR-2002
DEFINITION AGENCOURT_6706053 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5752141 5', mRNA sequence.
ACCESSION BM920874
VERSION BM920874.1 GI:19371253
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1045)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM12785 row: n column: 14
High quality sequence stop: 415.
FEATURES Location/Qualifiers
source 1. .1045
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5752141"
/clone_lib="NIH_MGC_115"
/lab_host="DH10B"
/note="Organ: pooled brain, lung, testis; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA source anonymous pool of 6 male brains, age range 23-27; 1 male lung, age 27; and 1 male testis, age 69. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.8 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 021. Note: this is a NIH_MGC Library."
BASE COUNT 176 a 343 c 341 g 183 t 2 others
ORIGIN

Query Match 48.7%; Score 671; DB 14; Length 1045;
Best Local Similarity 95.4%; Pred. No. 1.1e-82;
Matches 733; Conservative 0; Mismatches 31; Indels 4; Gaps 4;

```
Qy        171 GCCGGGAGGGGGGACTGCGAGAGGACCCCGGCGTCCGGGCTCCCGGTGCCAGCGCTATGA 230
              ||| ||      ||||||||||||||||||||||||||||||||||||||||||||||||||
Db         37 GCCTGGTCCACAGACTGCGAGAGGACCCCGGCGTCCGGGCTCCCGGTGCCAGCGCTATGA 96

Qy        231 GGCCACTCCTCGTCCTGCTGCTCCTGGGCCTGGCGGCCGGCTCGCCCCCACTGGACGACA 290
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db         97 GGCCACTCCTCGTCCTGCTGCTCCTGGGCCTGGCGGGCGGCTCGCCCCCACTGGACGACA 156

Qy        291 ACAAGATCCCCAGCCTCTGCCCGGGGCACCCCGGCCTTCCAGGCACGCCGGGCCACCATG 350
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        157 ACAAGATCCCCAGCCTCTGCCCGGGGCACCCCGGCCTTCCAGGCACGCCGGGCCACCATG 216

Qy        351 GCAGCCAGGGCTTGCCGGGCCGCGATGGCCGCGACGGCCGCGACGGCGCGCCCGGGGCTC 410
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        217 GCAGCCAGGGCTTGCCGGGCCGCGATGGCCGCGACGGCCGCGACGGCGCGCCCGGGGCTC 276

Qy        411 CGGGAGAGAAAGGCGAGGGCGGGAGGCCGGGACTGCCGGGACCTCGAGGGGACCCCGGGC 470
              ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        277 CGGGAGATAAAGGCGAGGGCGGGAGGCCGGGACTGCCGGGACCTCGAGGGGACCCCGGGC 336

Qy        471 CGCGAGGAGAGGCGGGACCCGCGGGGCCCACCGGGCCTGCCGGGGAGTGCTCGGTGCCTC 530
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        337 CGCGAGGAGAGGCGGGACCCGCGGGGCCCACCGGGCCTGCCGGGGAGTGCTCGGTGCCTC 396

Qy        531 CGCGATCCGCCTTCAGCGCCAAGCGCTCCGAGAGCCGGGTGCCTCCGCCGTCTGACGCAC 590
              |||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db        397 CGCGATCCGCCTTCAGCGCCAAGCGCTCCGAGAGCCGGGTGCCTCCGCCGACTGACGCAC 456

Qy        591 CCTTGCCCTTCGACCGCGTGCTGGTGAACGAGCAGGGACATTACGACGCCGTCACCGGCA 650
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        457 CCTTGCCCTTCGACCGCGTGCTGGTGAACGAGCAGGGACATTACGACGCCGTCACCGGCA 516

Qy        651 AGTTCACCTGCCAGGTGCCTGGGGTCTACTACTTCGCCGTCCATGCCACCGTCTACCGGG 710
              || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        517 AGATCACCTGCCAGGTGCCTGGGGTCTACTACTTCGCCGTCCATGCCACCGTCTACCGGG 576

Qy        711 CCAGCCTGCAGTTTGATCTGGTGAAGAATGGCGAATCCATTGCCTCTTTCTTCCAGTTTT 770
              |||||||||||||||||||| |||||||||||||||||||||||||||  |||| |||
Db        577 CCAGCCTGCAGTTTGATCTGGCGAAGAATGGCGAATCCATTGCCTCTTTGCTCCATGTTT 636

Qy        771 TCGGGGGGTGGCCCAAGCCAGCCTCGCTCTCGGGGGGGGCCATGGTGAGGCTGGAGCCTG 830
              |||| |||||||||| |||||||||  |||||||||||||||||||| ||||||||||||
Db        637 TCGGTGGGTGGCCCAAACCAGCCTCGATCTCGGGGGGGGCCATGGTGATGCTGGAGCCTG 696

Qy        831 AGGACCAAGTGTGGGTGCAGGTG-GGTGTGGGTGACTACA-TTGGCATCTATGCCA-GCA 887
              | |||||||||| ||||| ||| |||||| |||||||| |||  ||||||| ||  ||
Db        697 ATGACCAAGTGTGNGTGCATGTGNGGTGTGCGTGACTACATTTGTGATCTATGGCATTCA 756

Qy        888 TCAAGACAGACAGCACCTTCT-CCGGATTTCTGGTGTACTCCGACTGG 934
              || |||||||||| ||||||| |||||||||||||| ||||||||||
Db        757 TCCAGACAGACAGGACCTTCTCCCGGATTTCTGGTGAACTCCGACTGG 804
```

RESULT 8
BI770921
LOCUS BI770921 792 bp mRNA linear EST 25-SEP-2001
DEFINITION 603059737F1 NIH_MGC_122 Homo sapiens cDNA clone IMAGE:5209065 5',

```
            mRNA sequence.
ACCESSION   BI770921
VERSION     BI770921.1  GI:15762499
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 792)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Life Technologies, Inc.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM11524  row: j  column: 10
            High quality sequence stop: 782.
FEATURES             Location/Qualifiers
     source          1. .792
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5209065"
                     /clone_lib="NIH_MGC_122"
                     /lab_host="DH10B"
                     /note="Organ: pooled lung and spleen; Vector: pCMV-SPORT6;
                     Site_1: NotI; Site_2: EcoRV (destroyed); RNA source
                     anonymous pool of 24 week female lung, 16 week female
                     spleen, and 20-22 week male spleens.  Library is oligo-dT
                     primed and directionally cloned (EcoRV site is destroyed
                     upon  cloning).  Average insert size 1.4 kb, insert size
                     range 1-3 kb. Library is normalized and enriched for
                     full-length clones and was  constructed by C. Gruber
                     (Invitrogen).  Research Genetics tracking  code 026. Note:
                     this is a NIH_MGC Library."
BASE COUNT      142 a    232 c    237 g    180 t      1 others
ORIGIN
```

Query Match 48.0%; Score 660.4; DB 13; Length 792;
Best Local Similarity 96.9%; Pred. No. 3.5e-81;
Matches 758; Conservative 0; Mismatches 16; Indels 8; Gaps 8;

```
Qy        562 GAGCCGGGTGCCTCCGCCGTCTGACGCACCCTTGCCCTTCGACCGCGTGCTGGTGAACGA 621
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GAGCCGGGTGCCTCCGCCGTCTGACGCACCCTTGCCCTTCGACCGCGTGCTGGTGAACGA 60

Qy        622 GCAGGGACATTACGACGCCGTCACCGGCAAGTTCACCTGCCAGGTGCCTGGGGTCTACTA 681
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GCAGGGACATTACGACGCCGTCACCGGCAAGTTCACCTGCCAGGTGCCTGGGGTCTACTA 120

Qy        682 CTTCGCCGTCCATGCCACCGTCTACCGGGCCAGCCTGCAGTTTGATCTGGTGAAGAATGG 741
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CTTCGCCGTCCATGCCACCGTCTACCGGGCCAGCCTGCAGTTTGATCTGGTGAAGAATGG 180

Qy        742 CGAATCCATTGCCTCTTTCTTCCAGTTTTTCGGGGGGTGGCCCAAGCCAGCCTCGCTCTC 801
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CGAATCCATTGCCTCTTTCTTCCAGTTTTTCGGGGGGTGGCCCAAGCCAGCCTCGCTCTC 240

Qy        802 GGGGGGGGCCATGGTGAGGCTGGAGCCTGAGGACCAAGTGTGGGTGCAGGTGGGTGTGGG 861
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GGGGGGGGCCATGGTGAGGCTGGAGCCTGAGGACCAAGTGTGGGTGCAGGTGGGTGTGGG 300

Qy        862 TGACTACATTGGCATCTATGCCAGCATCAAGACAGACAGCACCTTCTCCGGATTTCTGGT 921
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TGACTACATTGGCATCTATGCCAGCATCAAGACAGACAGCACCTTCTCCGGATTTCTGGT 360

Qy        922 GTACTCCGACTGGCACAGCTCCCCAGTCTTTGCTTAGTGCCCACTGCAAAGTGAGCTCAT 981
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GTACTCCGACTGGCACAGCTCCCCAGTCTTTGCTTAGTGCCCACTGCAAAGTGAGCTCAT 420

Qy        982 GCTCTCACTCCTAGAAGGAGGGTGTGAGGCTGACAACCAGGTCATCCAGGAGGGC-TGGC 1040
              |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db        421 GCTCTCACTCCTAGAAGGAGGGTGTGAGGCTGACAACCAGGTCATCCAGGAGGGCTTGGC 480

Qy       1041 CCCCCTGGAATATTGTGAATGACTAGGGAGGTGGGGTAGAGC-ACTCTCCGTCCTGCTGC 1099
              |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db        481 CCCCCTGGAATATTGTGAATGACTAGGGAGGTGGGGTAGAGCAACTCTCCGTCCTGCTGC 540

Qy       1100 TGGCAAGGAATGGGAACAGTGGCTGTCTGCGATCAGGTCTGG-CAGCATGGGGCAGTGGC 1158
              |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db        541 TGGCAAGGAATGGGAACAGTGGCTGTCTGCGATCAGGTCTGGCCAGCATGGGGCAGTGGC 600

Qy       1159 TGGATTTCTGCCCAAGACCAG-AGGAGTGTGCTGTGCTGGCAAGTG-TAAGTCCCCCAGT 1216
              |||||||||||||||||||||| |||||||||||||||||||||| ||||||||||||||
Db        601 TGGATTTCTGCCCAAGACCAGCAGGAGTGTGCTGTGCTGGCAAGTGTTAAGTCCCCCAGT 660

Qy       1217 T-GCTCTGGTCCAGGAGCCCACG-GTGGGGTGCTCTCTT-CCTGGTCCTCTGCTTCTCTG 1273
              | ||||||||||||||||||||| |||| |||||||||| ||||||||||||||||||||
Db        661 TAGCTCTGGTCCAGGAGCCCACGAGTGGAGTGCTCTCTTACCTGGTCCTCTGCTTCTCTG 720

Qy       1274 GATCCTCCCCACCCCCTCCTGCTCCTGGGGCCGGCCCTTTTCTCAGAGATCACTCAATAA 1333
              |||||||||||| |||| |      |  ||| | | ||||||||||||||||||||||  ||
Db        721 GATCCTCCCCAACCCCCTCCTGCTCCTGGGGCCGGCCTTTTCTCAGAGATCACTCATAAA 780

Qy       1334 AC 1335
              ||
Db        781 AC 782
```

```
RESULT 9
BI490062
LOCUS       BI490062                887 bp    mRNA    linear   EST 28-AUG-2001
DEFINITION  603031867F1 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5172880 5',
            mRNA sequence.
ACCESSION   BI490062
VERSION     BI490062.1  GI:15329290
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 887)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Life Technologies, Inc.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM11430  row: f  column: 17
            High quality sequence stop: 815.
FEATURES             Location/Qualifiers
     source          1. .887
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5172880"
                     /clone_lib="NIH_MGC_115"
                     /lab_host="DH10B"
                     /note="Organ: pooled brain, lung, testis; Vector:
                     pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
                     source anonymous pool of 6 male brains, age range 23-27; 1
                     male lung, age 27; and 1 male testis, age 69.  Library is
                     oligo-dT primed and directionally cloned (EcoRV site is
                     destroyed upon  cloning).  Average insert size 1.8 kb,
                     insert size range 1-3 kb. Library is normalized and
                     enriched for full-length clones and was  constructed by C.
```

```
                Gruber (Invitrogen).  Research Genetics tracking  code
                021. Note: this is a NIH_MGC Library."
BASE COUNT      146 a    277 c    307 g    157 t
ORIGIN

  Query Match           46.6%;  Score 641.8;  DB 13;  Length 887;
  Best Local Similarity  84.7%;  Pred. No. 1.1e-78;
  Matches  858;  Conservative    0;  Mismatches   22;  Indels  133;  Gaps    7;

Qy         55 CAGGACTGGGGTGACGGCAGGGCAGGGGGCGCCTGGCCGGGGAGAAGCGCGGGGGCTGGA 114
               ||||||||||||||||||| || ||||||||||||||||||||||||||||||||||||
Db          1 CAGGACTGGGGTGACGGCAGGTCATGGGGCGCCTGGCCGGGGAGAAGCGCGGGGGCTGGA 60

Qy        115 GCACCACCAACTGGAGGGTCCGGAGTAGCGAGCGCCCCGAAGGAGGCCATCGGGGAGCCG 174
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GCACCACCAACTGGAGGGTCCGGAGTAGCGAGCGCCCCGAAGGAGGCCATCGGGGAGCCG 120

Qy        175 GGAGGGGGGACTGCGAGAGGACCCCGGCGTCCGGGCTCCCGGTGCCAGCGCTATGAGGCC 234
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GGAGGGGGGACTGCGAGAGGACCCCGGCGTCCGGGCTCCCGGTGCCAGCGCTATGAGGCC 180

Qy        235 ACTCCTCGTCCTGCTGCTCCTGGGCCTGGCGGCCGGCTCGCCCCCACTGGACGACAACAA 294
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ACTCCTCGTCCTGCTGCTCCTGGGCCTGGCGGCCGGCTCGCCCCCACTGGACGACAACAA 240

Qy        295 GATCCCCAGCCTCTGCCCGGGGCACCCCGGCCTTCCAGGCACGCCGGGCCACCATGGCAG 354
              ||||||||||||||||
Db        241 GATCCCCAGCCTCTGC---------------------------------------------- 256

Qy        355 CCAGGGCTTGCCGGGCCGCGATGGCCGCGACGGCCGCGACGGCGCGCCCGGGGCTCCGGG 414
Db        257 ---------------------------------------------------------- 256

Qy        415 AGAGAAAGGCGAGGGCGGGAGGCCGGGACTGCCGGGACCTCGAGGGGACCCCGGGCCGCG 474
Db        257 ----------------------CCGGGACTGCCGGGACCTCGAGGGGACCCCGGGCCGCG 294

Qy        475 AGGAGAGGCGGGACCCGCGGGGCCCACCGGGCCTGCCGGGGAGTGCTCGGTGCCTCCGCG 534
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        295 AGGAGAGGCGGGACCCGCGGGGCCCACCGGGCCTGCCGGGGAGTGCTCGGTGCCTCCGCG 354

Qy        535 ATCCGCCTTCAGCGCCAAGCGCTCCGAGAGCCGGGTGCCTCCGCCGTCTGACGCACCCTT 594
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        355 ATCCGCCTTCAGCGCCAAGCGCTCCGAGAGCCGGGTGCCTCCGCCGTCTGACGCACCCTT 414

Qy        595 GCCCTTCGACCGCGTGCTGGTGAACGAGCAGGGACATTACGACGCCGTCACCGGCAAGTT 654
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        415 GCCCTTCGACCGCGTGCTGGTGAACGAGCAGGGACATTACGACGCCGTCACCGGCAAGTT 474

Qy        655 CACCTGCCAGGTGCCTGGGGTCTACTACTTCGCCGTCCATGCCACCGTCTACCGGGCCAG 714
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        475 CACCTGCCAGGTGCCTGGGGTCTACTACTTCGCCGTCCATGCCACCGTCTACCGGGCCAG 534

Qy        715 CCTGCAGTTTGATCTGGTGAAGAATGGCGAATCCATTGCCTCTTTCTTCCAGTTTTTCGG 774
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        535 CCTGCAGTTTGATCTGGTGAAGAATGGCGAATCCATTGCCTCTTTCTTCCAGTTTTTCGG 594

Qy        775 GGGGTGGCCCAAGCCAGCCTCGCTCTCGGGGGGGGCCATGGTGAGGCTGGAGCCTGAGGA 834
              |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db        595 GGGGTGGCCCAAGCCAG-CTCGCTCTCGGGGGGGGCCATGGTGAGGCTGGAGCCTGAGGA 653

Qy        835 CCAAGTGTGGGTGCAGGTGGGTGTGGGTGACTACATTGGCATCTATGCCAGCATCAAGAC 894
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        654 CCAAGTGTGGGTGCAGGTGGGTGTGGGTGACTACATTGGCATCTATGCCAGCATCAAGAC 713

Qy        895 AGACAGCACCTTCTCCGGATTTC-TGGTGTACTCCGACT-GGCACAGCTCCCCAGTCTTT 952
              |||||||||||||||||||||| |||||||||||||||| |||||||||||     |||
Db        714 AGACAGCACCTTCTCCGGATTTCTTGGTGTACTCCGACTGGGCACAGCTCCCAGTCTTTT 773

Qy        953 GCTTAGTGCCCACTGCAAAGT--GAGCTCATGCTCTC-ACTCCTAGAAGGAGGGTGTGAG 1009
              ||||||||||||||||||||      ||||||||| | ||||||||||||||||||||||
Db        774 GCTTAGTGCCCACTGCAAAGTTGAGCTTCATGCTCTCAATTCCTAGAAGGAGGGTGTGAG 833

Qy       1010 GCTGACAACCAGG-TCATCCAGGAGGGCTGGCCCCCCTGGAATATTGTGAATG 1061
              |||||| |||| |||||||||| ||| | ||||||||||||||||   |||
Db        834 CGTGACAAACAGGTTCATCCAGGAAGGCGTGGCCCCCTGGAATATTGCTGATG 886


RESULT 10
BM926477
LOCUS       BM926477                1623 bp    mRNA    linear   EST 12-MAR-2002
DEFINITION  AGENCOURT_6644538 NIH_MGC_122 Homo sapiens cDNA clone IMAGE:5766864
            5', mRNA sequence.
ACCESSION   BM926477
VERSION     BM926477.1  GI:19376856
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1623)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Life Technologies, Inc.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM12824  row: d  column: 01
            High quality sequence stop: 558.
FEATURES             Location/Qualifiers
     source          1. .1623
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5766864"
                     /clone_lib="NIH_MGC_122"
                     /lab_host="DH10B"
                     /note="Organ: pooled lung and spleen; Vector: pCMV-SPORT6;
                     Site_1: NotI; Site_2: EcoRV (destroyed); RNA source
                     anonymous pool of 24 week female lung, 16 week female
                     spleen, and 20-22 week male spleens.  Library is oligo-dT
                     primed and directionally cloned (EcoRV site is destroyed
                     upon  cloning).  Average insert size 1.4 kb, insert size
                     range 1-3 kb. Library is normalized and enriched for
                     full-length clones and was  constructed by C. Gruber
                     (Invitrogen).  Research Genetics tracking  code 026. Note:
                     this is a NIH_MGC Library."
BASE COUNT      263 a    505 c    550 g    299 t      6 others
ORIGIN

  Query Match           45.8%;  Score 631;  DB 14;  Length 1623;
  Best Local Similarity  100.0%;  Pred. No. 2.4e-77;
  Matches  631;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         71 GCAGGGCAGGGGGCGCCTGGCCGGGGAGAAGCGCGGGGGCTGGAGCACCACCAACTGGAG 130
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GCAGGGCAGGGGGCGCCTGGCCGGGGAGAAGCGCGGGGGCTGGAGCACCACCAACTGGAG 60

Qy        131 GGTCCGGAGTAGCGAGCGCCCCGAAGGAGGCCATCGGGGAGCCGGGAGGGGGGACTGCGA 190
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GGTCCGGAGTAGCGAGCGCCCCGAAGGAGGCCATCGGGGAGCCGGGAGGGGGGACTGCGA 120

Qy        191 GAGGACCCCGGCGTCCGGGCTCCCGGTGCCAGCGCTATGAGGCCACTCCTCGTCCTGCTG 250
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GAGGACCCCGGCGTCCGGGCTCCCGGTGCCAGCGCTATGAGGCCACTCCTCGTCCTGCTG 180

Qy        251 CTCCTGGGCCTGGCGGCCGGCTCGCCCCCACTGGACGACAACAAGATCCCCAGCCTCTGC 310
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CTCCTGGGCCTGGCGGCCGGCTCGCCCCCACTGGACGACAACAAGATCCCCAGCCTCTGC 240

Qy        311 CCGGGGCACCCCGGCCTTCCAGGCACGCCGGGCCACCATGGCAGCCAGGGCTTGCCGGGC 370
```

```
Db        241 CCGGGGCACCCCGGCCTTCCAGGCACGCCGGGCCACCATGGCAGCCAGGGCTTGCCGGGC 300

Qy        371 CGCGATGGCCGCGACGGCCGCGACGGCGCGCCCGGGGCTCCGGGAGAGAAAGGCGAGGGC 430

Db        301 CGCGATGGCCGCGACGGCCGCGACGGCGCGCCCGGGGCTCCGGGAGAGAAAGGCGAGGGC 360

Qy        431 GGGAGGCCGGGACTGCCGGGACCTCGAGGGGACCCCGGGCCGCGAGGAGAGGCGGGACCC 490

Db        361 GGGAGGCCGGGACTGCCGGGACCTCGAGGGGACCCCGGGCCGCGAGGAGAGGCGGGACCC 420

Qy        491 GCGGGGCCCACCGGGCCTGCCGGGGAGTGCTCGGTGCCTCCGCGATCCGCCTTCAGCGCC 550

Db        421 GCGGGGCCCACCGGGCCTGCCGGGGAGTGCTCGGTGCCTCCGCGATCCGCCTTCAGCGCC 480

Qy        551 AAGCGCTCCGAGAGCCGGGTGCCTCCGCCGTCTGACGCACCCTTGCCCTTCGACCGCGTG 610

Db        481 AAGCGCTCCGAGAGCCGGGTGCCTCCGCCGTCTGACGCACCCTTGCCCTTCGACCGCGTG 540

Qy        611 CTGGTGAACGAGCAGGGACATTACGACGCCGTCACCGGCAAGTTCACCTGCCAGGTGCCT 670

Db        541 CTGGTGAACGAGCAGGGACATTACGACGCCGTCACCGGCAAGTTCACCTGCCAGGTGCCT 600

Qy        671 GGGGTCTACTACTTCGCCGTCCATGCCACCG 701

Db        601 GGGGTCTACTACTTCGCCGTCCATGCCACCG 631
```

```
RESULT 11
BG706609
LOCUS       BG706609                 630 bp    mRNA    linear   EST 07-MAY-2001
DEFINITION  602672994F1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:4795558 5',
            mRNA sequence.
ACCESSION   BG706609
VERSION     BG706609.1  GI:13982121
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 630)
  AUTHORS   NIH-MGC http://mgc.ncl.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
            cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
            Toshiyuki and Piero Carninci (RIKEN)
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM10678  row: 1  column: 23
            High quality sequence stop: 630.
FEATURES             Location/Qualifiers
     source          1. .630
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4795558"
                     /clone_lib="NIH_MGC_96"
                     /tissue_type="hypothalamus"
                     /lab_host="DH10B"
                     /note="Organ: brain; Vector: pBluescriptR (modified
                     pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag
                     ); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTTVN-3',
                     size-selected for  average insert size 2.3 kb and
                     normalized to ROT 5. This is a primary library enriched
                     for full-length  clones and constructed using the
                     Cap-trapper method (Carninci, in preparation).  Library
                     constructed by M. Brownstein (NIMH/NHGRI, National
                     Institutes of Health). Note: this is a NIH_MGC Library."
```

```
BASE COUNT       93 a    215 c    257 g     65 t
ORIGIN

  Query Match           44.1%;  Score 607.2;  DB 12;  Length 630;
  Best Local Similarity  99.5%;  Pred. No. 7.1e-74;
  Matches  609;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          7 TTCTCTTGGAGTCTGGGAGGAGGAAAGCGGAGCCGGCAGGGAGCGAACCAGGACTGGGGT 66

Db         19 TCCTCTTGGAGTCTGGGAGGAGGAAAGCGGAGCCGGCAGGGAGCGAACCAGGACTGGGGT 78

Qy         67 GACGGCAGGGCAGGGGGCGCCTGGCCGGGGAGAAGCGCGGGGGCTGGAGCACCACCAACT 126

Db         79 GACGGCAGGGCAGGGGGCGCCTGGCCGGGGAGAAGCGCGGGGGCTGGAGCACCACCAACT 138

Qy        127 GGAGGGTCCGGAGTAGCGAGCGCCCCGAAGGAGGCCATCGGGGAGCCGGGAGGGGGGACT 186

Db        139 GGAGGGTCCGGAGTAGCGAGCGCCCCGAAGGAGGCCATCGGGGAGCCGGGAGGGGGGACT 198

Qy        187 GCGAGAGGACCCCGGCGTCCGGGCTCCCGGTGCCAGCGCTATGAGGCCACTCCTCGTCCT 246

Db        199 GCGAGAGGACCCCGGCGTCCGGGCTCCCGGTGCCAGCGCTATGAGGCCACTCCTCGTCCT 258

Qy        247 GCTGCTCCTGGGCCTGGCGGCCGGCTCGCCCCCACTGGACGACAACAAGATCCCCAGCCT 306

Db        259 GCTGCTCCTGGGCCTCGCGGCCGGCTCGCCCCCACTGGACGACAACAAGATCCCCAGCCT 318

Qy        307 CTGCCCGGGGCACCCCGGCCTTCCAGGCACGCCGGGCCACCATGGCAGCCAGGGCTTGCC 366

Db        319 CTGCCCGGGGCACCCCGGCCTTCCAGGCACGCCGGGCCACCATGGCAGCCGGGGCTTGCC 378

Qy        367 GGGCCGCGATGGCCGCGACGGCCGCGACGGCGCGCCCGGGGCTCCGGGAGAGAAAGGCGA 426

Db        379 GGGCCGCGATGGCCGCGACGGCCGCGACGGCGCGCCCGGGGCTCCGGGAGAGAAAGGCGA 438

Qy        427 GGGCGGGAGGCCGGGACTGCCGGGACCTCGAGGGGACCCCGGGCCGCGAGGAGAGGCGGG 486

Db        439 GGGCGGGAGGCCGGGACTGCCGGGACCTCGAGGGGACCCCGGGCCGCGAGGAGAGGCGGG 498

Qy        487 ACCCGCGGGGCCCACCGGGCCTGCCGGGGAGTGCTCGGTGCCTCCGCGATCCGCCTTCAG 546

Db        499 ACCCGCGGGGCCCACCGGGCCTGCCGGGGAGTGCTCGGTGCCTCCGCGATCCGCCTTCAG 558

Qy        547 CGCCAAGCGCTCCGAGAGCCGGGTGCCTCCGCCGTCTGACGCACCCTTGCCCTTCGACCG 606

Db        559 CGCCAAGCGCTCCGAGAGCCGGGTGCCTCCGCCGTCTGACGCACCCTTGCCCTTCGACCG 618

Qy        607 CGTGCTGGTGAA 618

Db        619 CGTGCTGGTGAA 630
```

```
RESULT 12
BM893587/c
LOCUS       BM893587                 596 bp    mRNA    linear   EST 29-APR-2002
DEFINITION  ij26g10.x1 Melton Normalized Human Islet 4 N4-HIS 1 Homo sapiens
            cDNA clone IMAGE:6136122 3' similar to TR:Q9UFX4 Q9UFX4
            HYPOTHETICAL 22.8 KD PROTEIN ;, mRNA sequence.
ACCESSION   BM893587
VERSION     BM893587.1  GI:19349055
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 596)
  AUTHORS   Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K.,
            Lemishka,I., Scearce,M., Brestelli,J., Gradwohl,G., Clifton,S.,
            Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blistain,A.,
            Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J., Cardenas
            ,M., Gibbons,M., McCann,R., Cole,R., Tsagareishvili,R., Williams,T.
            , Jackson,Y. and Bowers,Y.
  TITLE     Endocrine Pancreas Consortium
```

```
  JOURNAL   Unpublished (2000)
  COMMENT   Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
            Endocrine Pancreas Consortium
            Harvard University, Howard Hughes Medical Institute
            Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
            MA 02138
            Tel: 617-495-1812
            Fax: 617-495-8557
            Email: dmelton@biohp.harvard.edu
            Library was constructed by Dr. Douglas Melton DNA sequencing by:
            Washington University Genome Sequencing Center For information on
            obtaining a clone please contact: Juliana Brown
            (brown@fas.harvard.edu)
            Seq primer: -40UP from Gibco
            High quality sequence stop: 441.
FEATURES             Location/Qualifiers
     source          1. .596
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6136122"
                     /clone_lib="Melton Normalized Human Islet 4 N4-HIS 1"
                     /sex="Both"
                     /tissue_type="Islets of Langerhans"
                     /dev_stage="Adult"
                     /lab_host="DH10B"
                     /note="Organ: Pancreas; Vector: pSPORT1; Site_1: Not 1;
                     Site_2: Sal 1; Starting library constructed using
                     SuperScript Plasmid Library kit (Life Technologies). cDNA
                     made by oligo-dT priming. Size-selected by column
                     fractionation; average insert size 1.08 kb. Library was
                     amplified once on solid support and plasmid DNA from
                     library was prepared. The library DNA was normalized by
                     method #4 from Bonaldo, Lennon, and Soares 1996 Genome
                     Research 6:791-806; 0.5 microgram single-stranded library
                     plasmid DNA was mixed with 5 micrograms PCR product
                     representing library inserts and hybridized to an Ecot of
                     20. Single-stranded (unhybridized) plasmids were isolated
                     by hydroxyapatite chromatography and used to make this
                     library."
BASE COUNT      139 a    183 c    169 g    105 t
ORIGIN

  Query Match           43.3%;  Score 596;  DB 14;  Length 596;
  Best Local Similarity  100.0%;  Pred. No. 2.5e-72;
  Matches  596;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        752 GCCTCTTTCTTCCAGTTTTTCGGGGGGTGGCCCAAGCCAGCCTCGCTCTCGGGGGGGGCC 811
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        596 GCCTCTTTCTTCCAGTTTTTCGGGGGGTGGCCCAAGCCAGCCTCGCTCTCGGGGGGGGCC 537

Qy        812 ATGGTGAGGCTGGAGCCTGAGGACCAAGTGTGGGTGCAGGTGGGTGTGGGTGACTACATT 871
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        536 ATGGTGAGGCTGGAGCCTGAGGACCAAGTGTGGGTGCAGGTGGGTGTGGGTGACTACATT 477

Qy        872 GGCATCTATGCCAGCATCAAGACAGACAGCACCTTCTCCGGATTTCTGGTGTACTCCGAC 931
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        476 GGCATCTATGCCAGCATCAAGACAGACAGCACCTTCTCCGGATTTCTGGTGTACTCCGAC 417

Qy        932 TGGCACAGCTCCCCAGTCTTTGCTTAGTGCCCACTGCAAAGTGAGCTCATGCTCTCACTC 991
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        416 TGGCACAGCTCCCCAGTCTTTGCTTAGTGCCCACTGCAAAGTGAGCTCATGCTCTCACTC 357

Qy        992 CTAGAAGGAGGGTGTGAGGCTGACAACCAGGTCATCCAGGAGGGCTGGCCCCCCCTGGAAT 1051
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        356 CTAGAAGGAGGGTGTGAGGCTGACAACCAGGTCATCCAGGAGGGCTGGCCCCCCCTGGAAT 297

Qy       1052 ATTGTGAATGACTAGGGAGGTGGGGTAGAGCACTCTCCGTCCTGCTGCTGGCAAGGAATG 1111
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        296 ATTGTGAATGACTAGGGAGGTGGGGTAGAGCACTCTCCGTCCTGCTGCTGGCAAGGAATG 237

Qy       1112 GGAACAGTGGCTGTCTGCGATCAGGTCTGGCAGCATGGGGCAGTGGCTGGATTTCTGCCC 1171
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        236 GGAACAGTGGCTGTCTGCGATCAGGTCTGGCAGCATGGGGCAGTGGCTGGATTTCTGCCC 177

Qy       1172 AAGACCAGAGGAGTGTGCTGTGCTGGCAAGTGTAAGTCCCCCAGTTGCTCTGGTCCAGGA 1231
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        176 AAGACCAGAGGAGTGTGCTGTGCTGGCAAGTGTAAGTCCCCCAGTTGCTCTGGTCCAGGA 117

Qy       1232 GCCCACGGTGGGGTGCTCTCTTCCTGGTCCTCTGCTTCTCTGGATCCTCCCCACCCCCTC 1291
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        116 GCCCACGGTGGGGTGCTCTCTTCCTGGTCCTCTGCTTCTCTGGATCCTCCCCACCCCCTC 57

Qy       1292 CTGCTCCTGGGGCCGGCCCTTTTCTCAGAGATCACTCAATAAACCTAAGAACCCTC 1347
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          56 CTGCTCCTGGGGCCGGCCCTTTTCTCAGAGATCACTCAATAAACCTAAGAACCCTC 1

RESULT 13
BM544255
LOCUS       BM544255                596 bp    mRNA    linear   EST 20-FEB-2002
DEFINITION  AGENCOURT_6490655 NIH_MGC_125 Homo sapiens cDNA clone IMAGE:5587759
            5', mRNA sequence.
ACCESSION   BM544255
VERSION     BM544255.1  GI:18775356
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 596)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Invitrogen
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM12357  row: m  column: 08
            High quality sequence stop: 595.
FEATURES             Location/Qualifiers
     source          1. .596
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5587759"
                     /clone_lib="NIH_MGC_125"
                     /lab_host="DH10B"
                     /note="Organ: ovary (pool of 3); Vector: pCMV-SPORT6;
                     Site_1: EcoRV (destroyed); Site_2: NotI; RNA source pool
                     of three ovaries, from females ranging in age from 38 to
                     49 yo. Library is oligo-dT primed and directionally cloned
                     (EcoRV site is destroyed upon cloning). Average insert
                     size 2.1 kb, insert size range 1-3.5 kb. Library is
                     normalized and enriched for full-length clones and was
                     constructed by C. Gruber (Invitrogen). Research Genetics
                     tracking code 036."
BASE COUNT      138 a    158 c    172 g    127 t      1 others
ORIGIN

  Query Match           42.6%;  Score 586.4;  DB 13;  Length 596;
  Best Local Similarity  99.7%;  Pred. No. 5e-71;
  Matches  587;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy        789 CAGCCTCGCTCTCGGGGGGGGCCATGGTGAGGCTGGAGCCTGAGGACCAAGTGTGGGTGC 848
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CAGCCTCGCTCTCGGGGGGGGCCATGGTGAGGCTGGAGCCTGAGGACCAAGTGTGGGTGC 60

Qy        849 AGGTGGGTGTGGGTGACTACATTGGCATCTATGCCAGCATCAAGACAGACAGCACCTTCT 908
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AGGTGGGTGTGGGTGACTACATTGGCATCTATGCCAGCATCAAGACAGACAGCACCTTCT 120

Qy        909 CCGGATTTCTGGTGTACTCCGACTGGCACAGCTCCCCAGTCTTTGCTTAGTGCCCACTGC 968
```

```
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CCGGATTTCTGGTGTACTCCGACTGGCACAGCTCCCCAGTCTTTGCTTAGTGCCCACTGC 180

Qy        969 AAAGTGAGCTCATGCTCTCACTCCTAGAAGGAGGGTGTGAGGCTGACAACCAGGTCATCC 1028
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AAAGTGAGCTCATGCTCTCACTCCTAGAAGGAGGGTGTGAGGCTGACAACCAGGTCATCC 240

Qy       1029 AGGAGGGCTGGCCCCCCCTGGAATATTGTGAATGACTAGGGAGGTGGGGTAGAGCACTCTC 1088
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AGGAGGGCTGGCCCCCCTGGAATATTGTGAATGACTAGGGAGGTGGGGTAGAGCACTCTC 300

Qy       1089 CGTCCTGCTGCTGGCAAGGAATGGGAACAGTGGCTGTCTGCGATCAGGTCTGGCAGCATG 1148
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CGTCCTGCTGCTGGCAAGGAATGGGAACAGTGGCTGTCTGCGATCAGGTCTGGCAGCATG 360

Qy       1149 GGGCAGTGGCTGGATTTCTGCCCAAGACCAGAGGAGTGTGCTGTGCTGGCAAGTGTAAGT 1208
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GGGCAGTGGCTGGATTTCTGCCCAAGACCAGAGGAGTGTGCTGTGCTGGCAAGTGTAAGT 420

Qy       1209 CCCCCAGTTGCTCTGGTCCAGGAGCCCACGGTGGGGTGCTCTCTTCCTGGTCCTCTGCTT 1268
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CCCCCAGTTGCTCTGGTCCAGGAGCCCACGGTGGGGTGCTCTCTTCCTGGTCCTCTGCTT 480

Qy       1269 CTCTGGATCCTCCCCACCCCCTCCTGCTCCTGGGGCCGGCCCTTTTCTCAGAGATCACTC 1328
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CTCTGGATCCTCCCCACCCCCTCCTGCTCCTGGGGCCGGCCCTTTTCTCAGAGATCACTC 540

Qy       1329 AATAAACCTAAGAACCCTCATAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1377
                 |||||||||||||||||||| ||||| ||||||||||||||||||||
Db        541 AATAAACCTAAGAACCCTCAAAAAAANAAAAAAAAAAAAAAAAAAAAAA 589


RESULT 14
BI821899
LOCUS       BI821899                 941 bp    mRNA    linear   EST 04-OCT-2001
DEFINITION  603035796F1 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5176967 5',
            mRNA sequence.
ACCESSION   BI821899
VERSION     BI821899.1  GI:15933449
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 941)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Life Technologies, Inc.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM11440  row: p  column: 24
            High quality sequence stop: 779.
FEATURES             Location/Qualifiers
     source          1. .941
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5176967"
                     /clone_lib="NIH_MGC_115"
                     /lab_host="DH10B"
                     /note="Organ: pooled brain, lung, testis; Vector:
                     pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
                     source anonymous pool of 6 male brains, age range 23-27; 1
                     male lung, age 27; and 1 male testis, age 69.  Library is
                     oligo-dT primed and directionally cloned (EcoRV site is
                     destroyed upon  cloning).  Average insert size 1.8 kb,
                     insert size range 1-3 kb. Library is normalized and
                     enriched for full-length clones and was  constructed by C.
                     Gruber (Invitrogen).  Research Genetics tracking  code
                     021. Note: this is a NIH_MGC Library."
BASE COUNT      130 a    320 c    329 g    162 t
ORIGIN

  Query Match           42.4%;  Score 583.2;  DB 13;  Length 941;
  Best Local Similarity   94.3%;  Pred. No. 1e-70;
  Matches  758;  Conservative    0;  Mismatches   28;  Indels   18;  Gaps   14;

Qy        191 GAGGACCCCGGCGTCCGGGCTCCCGGTGCCAGCGCTATGAG-GCCACTCCTCGTCCTGCT 249
              |||||||||||||||||||||||||||||||||||||||| |||||||||||||||||||
Db        136 GAGGACCCCGGCGTCCGGGCTCCCGGTGCCAGCGCTATGAGTGCCACTCCTCGTCCTGCT 195

Qy        250 GCTCCTGGGCCTGGCGGCCGGCTCGCCCCCACTGGACGACAACAAGATCCCCAGCCTCTG 309
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        196 GCTCCTGGGCCTGGCGGCCGGCTCGCCCCCACTGGACGACAACAAGATCCCCAGCCTCTG 255

Qy        310 CCCGGGGCACCCCGGCCTTCCAGGCACGCCGGGCCACCATGGCAGCCA-GGGCTTGCCGG 368
              |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db        256 CCCGGGGCACCCCGGCCTTCCAGGCACGCCGGGCCACCATGGCAGCCATGGGCTTGCCGG 315

Qy        369 GCCGCGATGGCCGCGACGGCCGCGACGGCGCGCCCGG-GGCTCCGGGAGAGAAAGGCGAG 427
              |||||||||||||||||||||||||||||||||||| |||||||||||||||||||||||
Db        316 GCCGCGATGGCCGCGACGGCCGCGACGGCGCGCCCGGTGGCTCCGGGAGAGAAAGGCGAG 375

Qy        428 GGCGGGAGGCCGGGACTGCCGGGACCTCGAGGGGACCCCGGGCCGCGAGGAGAGGCGGGA 487
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        376 GGCGGGAGGCCGGGACTGCCGGGACCTCGAGGGGACCCCGGGCCGCGAGGAGAGGCGGGA 435

Qy        488 CCCGCGGGGCCCACCGGG--CCTGCCGGGGAGTGCTCGGTGCCTCCGCGATCCGCCTTCA 545
              |||||||||||||||||  |   | ||||||||||||||||||||||||||||| |||||
Db        436 CCCGCGGGGCCCACCGGGGCCTGCCGGGGGAGTGCTCGGTGCCTCCGCGATCCGGCTTCA 495

Qy        546 GCGCCAAGCGCTCCGAGAGCC--GGGTGCCTCCGCCGTCTGACGCACCCTTG-CCCTTCG 602
              |||||||||||||||||||||  || |||||||||||||||||||||||||| |||||||
Db        496 GCGCCAAGCGCTCCGAGAGCCTGGGTTGCCTCCGCCGTCTGACGCACCCTTGCCCCTTCG 555

Qy        603 ACCGCGTGCTGGTGAACGAGCAGGGACATTACGACGCCGTCACCGGCAAGTTCACCTGCC 662
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        556 ACCGCGTGCTGGTGAACGAGCAGGGACATTACGACGCCGTCACCGGCAAGTTCACCTGCC 615

Qy        663 -AGGTGCCTGGGGTCTACTACTTCGCCGTCCA-TGCCACCGTCTACCGGGCCAGCCTGCA 720
               |||||||||||||||||||||||||||||| |||||||||||||||||||||||||||
Db        616 GAGGTGCCTGGGGTCTACTACTTCGCCGTCCATTGCCACCGTCTACCGGGCCAGCCTGCA 675

Qy        721 GTTTGATCTGGTGAAGAATGGCGAATCCATTGCCTCTTTCTTCCAGTTTTTCGGGGGGTG 780
              |||||||||||||||||||||||||||||||||||||||||||||||| ||||||||  |||
Db        676 GTTTGATCTGGTGAAGAATGGCGAATCCATTGCCTCTTTCTTCCAG-TTTTCGGGCAGTG 734

Qy        781 GCCCAAGCCAGCCTCGCTCTCGGGGGGGGCCATGGTGAGGCTGGAG-CCTGAGGACCAAG 839
              |||||||||||||||||||||||| |||||||||||||||||| ||||||||||||||||
Db        735 GCCCAAGCCAGCCTCGCTCTCGGGGGGCGCCATGGTGAGGCTGGAGCCCTGAGGACCAAG 794

Qy        840 TGTGGGTGCAGGT-GGGTGTGGGTGACTACATTGGCATCTATGCCAGCATCAAGACAGAC 898
              ||||||| | ||| |||||||||||||||| ||||| |||||||||||||| || |||||
Db        795 TGTGGGTTCCGGTGGGGTGTGGGTGACTACTTTGGCTTCTATGCCAGCATC-AGGCAGAC 853

Qy        899 AGCACCTTCTCCGGATTTCTGGTGTACTCCGACTGGCACAGCTCCCCAGTCTTTGCTTAG 958
              || |||||||||||| ||||||||||||||||||||||||||  |||  |||||  |  |
Db        854 AGGACCTTCTCCGGA-TTCTGGTGTACTCCGACTGGCACAGTCCCCAGGTCTTGCTTGTG 912

Qy        959 TGCCCACTGCAAAGTGAGCTCATG 982
                 || |||  |||||||||| ||
Db        913 ---CCCCTGTCAAGTGAGCTCCTG 933


RESULT 15
BM893691/c
LOCUS       BM893691                 578 bp    mRNA    linear   EST 29-APR-2002
DEFINITION  1j28e08.x1 Melton Normalized Human Islet 4 N4-HIS 1 Homo sapiens
            cDNA clone IMAGE:6136046 3' similar to TR:Q9UFX4 Q9UFX4
            HYPOTHETICAL 22.8 KD PROTEIN ;, mRNA sequence.
```

```
ACCESSION   BM893691
VERSION     BM893691.1  GI:19349159
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 578)
  AUTHORS   Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K.,
            Lemishka,I., Scearce,M., Brestelli,J., Gradwohl,G., Clifton,S.,
            Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blistain,A.,
            Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J., Cardenas
            ,M., Gibbons,M., McCann,R., Cole,R., Tsagareishvili,R., Williams,T.
            , Jackson,Y. and Bowers,Y.
  TITLE     Endocrine Pancreas Consortium
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
            Endocrine Pancreas Consortium
            Harvard University, Howard Hughes Medical Institute
            Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
            MA 02138
            Tel: 617-495-1812
            Fax: 617-495-8557
            Email: dmelton@biohp.harvard.edu
            Library was constructed by Dr. Douglas Melton DNA sequencing by:
            Washington University Genome Sequencing Center For information on
            obtaining a clone please contact: Juliana Brown
            (brown@fas.harvard.edu)
            Seq primer: -40UP from Gibco
            High quality sequence stop: 408.
FEATURES             Location/Qualifiers
     source          1. .578
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6136046"
                     /clone_lib="Melton Normalized Human Islet 4 N4-HIS 1"
                     /sex="Both"
                     /tissue_type="Islets of Langerhans"
                     /dev_stage="Adult"
                     /lab_host="DH10B"
                     /note="Organ: Pancreas; Vector: pSPORT1; Site_1: Not 1;
                     Site_2: Sal 1; Starting library constructed using
                     SuperScript Plasmid Library kit  (Life Technologies). cDNA
                     made by oligo-dT priming. Size-selected by  column
                     fractionation; average insert size 1.08 kb. Library was
                     amplified  once on solid support and plasmid DNA from
                     library was prepared. The  library DNA was normalized by
                     method #4 from Bonaldo, Lennon, and Soares  1996 Genome
                     Research 6:791-806; 0.5 microgram single-stranded library
                     plasmid DNA was mixed with 5 micrograms PCR product
                     representing library  inserts and hybridized to an Ecot of
                     20. Single-stranded (unhybridized)  plasmids were isolated
                     by hydroxyapatite chromatography and used to make  this
                     library."
BASE COUNT      130 a    181 c    163 g    104 t
ORIGIN
```

Query Match 42.0%; Score 578; DB 14; Length 578;
Best Local Similarity 100.0%; Pred. No. 7.1e-70;
Matches 578; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy        770 TTCGGGGGGTGGCCCAAGCCAGCCTCGCTCTCGGGGGGGGCCATGGTGAGGCTGGAGCCT 829
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        578 TTCGGGGGGTGGCCCAAGCCAGCCTCGCTCTCGGGGGGGGCCATGGTGAGGCTGGAGCCT 519

Qy        830 GAGGACCAAGTGTGGGTGCAGGTGGGTGTGGGTGACTACATTGGCATCTATGCCAGCATC 889
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        518 GAGGACCAAGTGTGGGTGCAGGTGGGTGTGGGTGACTACATTGGCATCTATGCCAGCATC 459

Qy        890 AAGACAGACAGCACCTTCTCCGGATTTCTGGTGTACTCCGACTGGCACAGCTCCCCAGTC 949
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        458 AAGACAGACAGCACCTTCTCCGGATTTCTGGTGTACTCCGACTGGCACAGCTCCCCAGTC 399

Qy        950 TTTGCTTAGTGCCCACTGCAAAGTGAGCTCATGCTCTCACTCCTAGAAGGAGGGTGTGAG 1009
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        398 TTTGCTTAGTGCCCACTGCAAAGTGAGCTCATGCTCTCACTCCTAGAAGGAGGGTGTGAG 339

Qy       1010 GCTGACAACCAGGTCATCCAGGAGGGCTGGCCCCCCCTGGAATATTGTGAATGACTAGGGA 1069
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        338 GCTGACAACCAGGTCATCCAGGAGGGCTGGCCCCCCCTGGAATATTGTGAATGACTAGGGA 279

Qy       1070 GGTGGGGTAGAGCACTCTCCGTCCTGCTGCTGGCAAGGAATGGGAACAGTGGCTGTCTGC 1129
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        278 GGTGGGGTAGAGCACTCTCCGTCCTGCTGCTGGCAAGGAATGGGAACAGTGGCTGTCTGC 219

Qy       1130 GATCAGGTCTGGCAGCATGGGGCAGTGGCTGGATTTCTGCCCAAGACCAGAGGAGTGTGC 1189
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        218 GATCAGGTCTGGCAGCATGGGGCAGTGGCTGGATTTCTGCCCAAGACCAGAGGAGTGTGC 159

Qy       1190 TGTGCTGGCAAGTGTAAGTCCCCCAGTTGCTCTGGTCCAGGAGCCCACGGTGGGGTGCTC 1249
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        158 TGTGCTGGCAAGTGTAAGTCCCCCAGTTGCTCTGGTCCAGGAGCCCACGGTGGGGTGCTC 99

Qy       1250 TCTTCCTGGTCCTCTGCTTCTCTGGATCCTCCCCACCCCCTCCTGCTCCTGGGGCCGGCC 1309
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         98 TCTTCCTGGTCCTCTGCTTCTCTGGATCCTCCCCACCCCCTCCTGCTCCTGGGGCCGGCC 39

Qy       1310 CTTTTCTCAGAGATCACTCAATAAACCTAAGAACCCTC 1347
              |||||||||||||||||||||||||||||||||||||
Db         38 CTTTTCTCAGAGATCACTCAATAAACCTAAGAACCCTC 1
```

Search completed: June 22, 2003, 05:15:25
Job time : 1908 secs

GenCore version 5.1.6


OM nucleic - nucleic search, using sw model

Run on: June 22, 2003, 03:34:16 ; Search time 94 Seconds
(without alignments)
4492.488 Million cell updates/sec

Title: US-09-943-851A-41
Perfect score: 1377
Sequence: 1 gactagttctcttggagtct.....aaaaaaaaaaaaaaaaaaa 1377

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 1322.2 | 96.0 | 1347 | 4 | US-09-140-804-1 | Sequence 1, Appli |
| 2 | 1296.6 | 94.2 | 1338 | 4 | US-09-336-536-1 | Sequence 1, Appli |
| 3 | 717 | 52.1 | 728 | 4 | US-09-336-536-2 | Sequence 2, Appli |
| 4 | 692.8 | 50.3 | 1263 | 4 | US-09-336-536-8 | Sequence 8, Appli |
| 5 | 606 | 44.0 | 1001 | 3 | US-09-188-930-218 | Sequence 218, App |
| 6 | 605 | 43.9 | 1015 | 3 | US-09-188-930-30 | Sequence 30, Appl |
| 7 | 564.2 | 41.0 | 729 | 4 | US-09-336-536-9 | Sequence 9, Appli |
| 8 | 504 | 36.6 | 729 | 4 | US-09-140-804-10 | Sequence 10, Appl |
| 9 | 257.8 | 18.7 | 601 | 4 | US-09-336-536-74 | Sequence 74, Appl |
| 10 | 149.6 | 10.9 | 393 | 3 | US-09-188-930-26 | Sequence 26, Appl |
| 11 | 90.6 | 6.6 | 1276 | 2 | US-08-463-911-1 | Sequence 1, Appli |
| 12 | 82.2 | 6.0 | 1313 | 2 | US-08-463-911-6 | Sequence 6, Appli |
| 13 | 82.2 | 6.0 | 4517 | 4 | US-09-140-804-9 | Sequence 9, Appli |
| 14 | 77.4 | 5.6 | 843 | 4 | US-09-118-408-23 | Sequence 23, Appl |
| 15 | 77.4 | 5.6 | 843 | 4 | US-09-506-855-23 | Sequence 23, Appl |
| 16 | 73 | 5.3 | 144 | 4 | US-09-140-804-28 | Sequence 28, Appl |
| 17 | 72 | 5.2 | 144 | 4 | US-09-140-804-29 | Sequence 29, Appl |
| 18 | 72 | 5.2 | 144 | 4 | US-09-140-804-31 | Sequence 31, Appl |
| 19 | 71.8 | 5.2 | 147 | 4 | US-09-140-804-30 | Sequence 30, Appl |
| 20 | 68.4 | 5.0 | 4411529 | 4 | US-09-103-840A-1 | Sequence 1, Appli |
| 21 | 65.2 | 4.7 | 2543 | 1 | US-08-555-669-11 | Sequence 11, Appl |
| 22 | 65.2 | 4.7 | 2543 | 3 | US-09-073-663-11 | Sequence 11, Appl |
| 23 | 62 | 4.5 | 985 | 4 | US-09-056-556-182 | Sequence 182, App |
| 24 | 62 | 4.5 | 985 | 4 | US-09-072-596-177 | Sequence 177, App |
| 25 | 62 | 4.5 | 4403765 | 4 | US-09-103-840A-2 | Sequence 2, Appli |
| 26 | 60.8 | 4.4 | 756 | 1 | US-08-642-255-50 | Sequence 50, Appl |
| 27 | 60.8 | 4.4 | 1839 | 1 | US-08-383-744-1 | Sequence 1, Appli |
| 28 | 60.8 | 4.4 | 1839 | 2 | US-08-999-336-1 | Sequence 1, Appli |
| 29 | 60.8 | 4.4 | 1839 | 5 | PCT-US96-01427-1 | Sequence 1, Appli |
| 30 | 59.8 | 4.3 | 1333 | 4 | US-09-227-357-51 | Sequence 51, Appl |
| 31 | 59.6 | 4.3 | 801 | 2 | US-08-770-379-16 | Sequence 16, Appl |
| 32 | 59.6 | 4.3 | 801 | 4 | US-08-757-669A-16 | Sequence 16, Appl |
| 33 | 59.6 | 4.3 | 801 | 4 | US-09-230-371A-16 | Sequence 16, Appl |
| 34 | 59.2 | 4.3 | 432 | 1 | US-08-642-255-48 | Sequence 48, Appl |
| 35 | 58.2 | 4.2 | 2769 | 4 | US-09-118-408-1 | Sequence 1, Appli |
| 36 | 58.2 | 4.2 | 2769 | 4 | US-09-506-855-1 | Sequence 1, Appli |
| 37 | 58 | 4.2 | 3181 | 1 | US-08-655-086-1 | Sequence 1, Appli |
| c 38 | 57.4 | 4.2 | 4411529 | 4 | US-09-103-840A-1 | Sequence 1, Appli |
| c 39 | 56.6 | 4.1 | 4403765 | 4 | US-09-103-840A-2 | Sequence 2, Appli |
| 40 | 56.4 | 4.1 | 444 | 4 | US-09-397-787-161 | Sequence 161, App |
| 41 | 56.4 | 4.1 | 4488 | 4 | US-08-406-030A-3 | Sequence 3, Appli |
| c 42 | 56 | 4.1 | 4257 | 2 | US-08-690-473-1 | Sequence 1, Appli |
| c 43 | 56 | 4.1 | 4257 | 4 | US-09-259-821A-1 | Sequence 1, Appli |
| c 44 | 56 | 4.1 | 4257 | 4 | US-08-843-659-1 | Sequence 1, Appli |
| 45 | 54.8 | 4.0 | 801 | 4 | US-09-298-568-3 | Sequence 3, Appli |

ALIGNMENTS

RESULT 1
US-09-140-804-1
; Sequence 1, Application US/09140804
; Patent No. 6197930
; GENERAL INFORMATION:
; APPLICANT: Sheppard, Paul O.
; APPLICANT: Humes, Jacqueline M.
; TITLE OF INVENTION: ADIPOCYTE-SPECIFIC PROTEIN HOMOLOGS
; FILE REFERENCE: 97-49
; CURRENT APPLICATION NUMBER: US/09/140,804
; CURRENT FILING DATE: 1998-08-26
; EARLIER APPLICATION NUMBER: 60/056,983
; EARLIER FILING DATE: 1997-08-26
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 1347
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (198)...(926)
US-09-140-804-1

Query Match 96.0%; Score 1322.2; DB 4; Length 1347;
Best Local Similarity 99.8%; Pred. No. 5.7e-257;
Matches 1324; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy      40 CGGCAGGGAGCGAACCAGGACTGGGGTGACGGCAGGGCAGGGGGCGCCTGGCCGGGGAGA 99
           ||  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      11 CGAGAGGGAGCGAACCAGGACTGGGGTGACGGCAGGGCAGGGGGCGCCTGGCCGGGGAGA 70

Qy     100 AGCGCGGGGGCTGGAGCACCACCAACTGGAGGGTCCGGAGTAGCGAGCGCCCCGAAGGAG 159
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      71 AGCGCGGGGGCTGGAGCACCACCAACTGGAGGGTCCGGAGTAGCGAGCGCCCCGAAGGAG 130

Qy     160 GCCATCGGGGAGCCGGGAGGGGGGACTGCGAGAGGACCCCGGCGTCCGGGCTCCCGGTGC 219
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     131 GCCATCGGGGAGCCGGGAGGGGGGACTGCGAGAGGACCCCGGCGTCCGGGCTCCCGGTGC 190

Qy     220 CAGCGCTATGAGGCCACTCCTCGTCCTGCTGCTCCTGGGCCTGGCGGCCGGCTCGCCCCC 279
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     191 CAGCGCTATGAGGCCACTCCTCGTCCTGCTGCTCCTGGGCCTGGCGGCCGGCTCGCCCCC 250

Qy     280 ACTGGACGACAACAAGATCCCCAGCCTCTGCCCGGGGCACCCCGGCCTTCCAGGCACGCC 339
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     251 ACTGGACGACAACAAGATCCCCAGCCTCTGCCCGGGGCACCCCGGCCTTCCAGGCACGCC 310

Qy     340 GGGCCACCATGGCAGCCAGGGCTTGCCGGGCCGCGATGGCCGCGACGGCCGCGACGGCGC 399
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        311 GGGCCACCATGGCAGCCAGGGCTTGCCGGGCCGCGATGGCCGCGACGGCCGCGACGGCGC 370
Qy        400 GCCCGGGGCTCCGGGAGAGAAAGGCGAGGGCGGGAGGCCGGGACTGCCGGGACCTCGAGG 459
Db        371 GCCCGGGGCTCCGGGAGAGAAAGGCGAGGGCGGGAGGCCGGGACTGCCGGGACCTCGAGG 430
Qy        460 GGACCCCGGGCCGCGAGGAGAGGCGGGACCCGCGGGGCCCACCGGGCCTGCCGGGGAGTG 519
Db        431 GGACCCCGGGCCGCGAGGAGAGGCGGGACCCGCGGGGCCCACCGGGCCTGCCGGGGAGTG 490
Qy        520 CTCGGTGCCTCCGCGATCCGCCTTCAGCGCCAAGCGCTCCGAGAGCCGGGTGCCTCCGCC 579
Db        491 CTCGGTGCCTCCGCGATCCGCCTTCAGCGCCAAGCGCTCCGAGAGCCGGGTGCCTCCGCC 550
Qy        580 GTCTGACGCACCCTTGCCCTTCGACCGCGTGCTGGTGAACGAGCAGGGACATTACGACGC 639
Db        551 GTCTGACGCACCCTTGCCCTTCGACCGCGTGCTGGTGAACGAGCAGGGACATTACGACGC 610
Qy        640 CGTCACCGGCAAGTTCACCTGCCAGGTGCCTGGGGTCTACTACTTCGCCGTCCATGCCAC 699
Db        611 CGTCACCGGCAAGTTCACCTGCCAGGTGCCTGGGGTCTACTACTTCGCCGTCCATGCCAC 670
Qy        700 CGTCTACCGGGCCAGCCTGCAGTTTGATCTGGTGAAGAATGGCGAATCCATTGCCTCTTT 759
Db        671 CGTCTACCGGGCCAGCCTGCAGTTTGATCTGGTGAAGAATGGCGAATCCATTGCCTCTTT 730
Qy        760 CTTCCAGTTTTTCGGGGGGTGGCCCAAGCCAGCCTCGCTCTCGGGGGGGGCCATGGTGAG 819
Db        731 CTTCCAGTTTTTCGGGGGGTGGCCCAAGCCAGCCTCGCTCTCGGGGGGGGCCATGGTGAG 790
Qy        820 GCTGGAGCCTGAGGACCAAGTGTGGGTGCAGGTGGGTGTGGGTGACTACATTGGCATCTA 879
Db        791 GCTGGAGCCTGAGGACCAAGTGTGGGTGCAGGTGGGTGTGGGTGACTACATTGGCATCTA 850
Qy        880 TGCCAGCATCAAGACAGACAGCACCTTCTCCGGATTTCTGGTGTACTCCGACTGGCACAG 939
Db        851 TGCCAGCATCAAGACAGACAGCACCTTCTCCGGATTTCTGGTGTACTCCGACTGGCACAG 910
Qy        940 CTCCCCAGTCTTTGCTTAGTGCCCACTGCAAAGTGAGCTCATGCTCTCACTCCTAGAAGG 999
Db        911 CTCCCCAGTCTTTGCTTAGTGCCCACTGCAAAGTGAGCTCATGCTCTCACTCCTAGAAGG 970
Qy       1000 AGGGTGTGAGGCTGACAACCAGGTCATCCAGGAGGGCTGGCCCCCCTGGAATATTGTGAA 1059
Db        971 AGGGTGTGAGGCTGACAACCAGGTCATCCAGGAGGGCTGGCCCCCCTGGAATATTGTGAA 1030
Qy       1060 TGACTAGGGAGGTGGGGTAGAGCACTCTCCGTCCTGCTGCTGGCAAGGAATGGGAACAGT 1119
Db       1031 TGACTAGGGAGGTGGGGTAGAGCACTCTCCGTCCTGCTGCTGGCAAGGAATGGGAACAGT 1090
Qy       1120 GGCTGTCTGCGATCAGGTCTGGCAGCATGGGGCAGTGGCTGGATTTCTGCCCAAGACCAG 1179
Db       1091 GGCTGTCTGCGATCAGGTCTGGCAGCATGGGGCAGTGGCTGGATTTCTGCCCAAGACCAG 1150
Qy       1180 AGGAGTGTGCTGTGCTGGCAAGTGTAAGTCCCCCAGTTGCTCTGGTCCAGGAGCCCACGG 1239
Db       1151 AGGAGTGTGCTGTGCTGGCAAGTGTAAGTCCCCCAGTTGCTCTGGTCCAGGAGCCCACGG 1210
Qy       1240 TGGGGTGCTCTCTTCCTGGTCCTCTGCTTCTCTGGATCCTCCCCACCCCCTCCTGCTCCT 1299
Db       1211 TGGGGTGCTCTCTTCCTGGTCCTCTGCTTCTCTGGATCCTCCCCACCCCCTCCTGCTCCT 1270
Qy       1300 GGGGCCGGCCCTTTTCTCAGAGATCACTCAATAAACCTAAGAACCCTCATAAAAAAAAAAA 1359
Db       1271 GGGGCCGGCCCTTTTCTCAGAGATCACTCAATAAACCTAAGAACCCTCAAAAAAAAAAAA 1330
Qy       1360 AAAAAAA 1366
Db       1331 AAAAAAA 1337

RESULT 2
US-09-336-536-1

; Sequence 1, Application US/09336536
; Patent No. 6406884
; GENERAL INFORMATION:
;  APPLICANT: Leiby, K.
;  APPLICANT: McKay, C.
;  APPLICANT: Bossone, S.
;  TITLE OF INVENTION: SECRETED PROTEINS AND USES THEREOF
;  FILE REFERENCE: 7853-144
;  CURRENT APPLICATION NUMBER: US/09/336,536
;  CURRENT FILING DATE: 1999-06-18
;  NUMBER OF SEQ ID NOS: 75
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
;   LENGTH: 1338
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-336-536-1

  Query Match           94.2%;  Score 1296.6;  DB 4;  Length 1338;
  Best Local Similarity  99.6%;  Pred. No. 7.7e-252;
  Matches 1310;  Conservative    0;  Mismatches    4;  Indels    1;  Gaps    1;

Qy         54 CCAGGACTGGGGTGACGGCAGGGCAGGGGGCGCCTGGCCGGGGAGAAGCGCGGGGGCTGG 113
Db         15 CCGGGACTGGGGTGACGGCAGGGCAGGGGGCGCCTGGCCGGGGAGAAGCGCGGGGGCTGG 74
Qy        114 AGCACCACCAACTGGAGGGTCCGGAGTAGCGAGCGCCCCGAAGGAGGCCATCGGGGAGCC 173
Db         75 AGCACCACCAACTGGAGGGTCCGGAGTAGCGAGCGCCCCGAAGGAGGCCATCGGGGAGCC 134
Qy        174 GGGAGGGGGGACTGCGAGAGGACCCCGGCGTCCGGGCTCCCGGTGCCAGCGCTATGAGGC 233
Db        135 GGGAGGGGGGACTGCGAGAGGACCCCGGCGTCCGGGCTCCCGGTGCCAGCGCTATGAGGC 194
Qy        234 CACTCCTCGTCCTGCTGCTCCTGGGCCTGGCGGCCGGCTCGCCCCCACTGGACGACAACA 293
Db        195 CACTCCTCGTCCTGCTGCTCCTGGGCCTGGCGGCCGGCTCGCCCCCACTGGACGACAACA 254
Qy        294 AGATCCCCAGCCTCTGCCCGGGGCACCCCGGCCTTCCAGGCACGCCGGGCCACCATGGCA 353
Db        255 AGATCCCCAGCCTCTGCCCGGGGCACCCCGGCCTTCCAGGCACGCCGGGCCACCATGGCA 314
Qy        354 GCCAGGGCTTGCCGGGCCGCGATGGCCGCGACGGCCGCGACGGCGCGCCCGGGGCTCCGG 413
Db        315 GCCAGGGCTTGCCGGGCCGCGATGGCCGCGACGGCCGCGACGGCGCGCCCGGGGCTCCGG 374
Qy        414 GAGAGAAAGGCGAGGGCGGGAGGCCGGGACTGCCGGGACCTCGAGGGGACCCCGGGCCGC 473
Db        375 GAGAGAAAGGCGAGGGCGGGAGG-CGGGACTGCCGGGACCTCGAGGGGACCCCGGGCCGC 433
Qy        474 GAGGAGAGGCGGGACCCGCGGGGCCCACCGGGCCTGCCGGGGAGTGCTCGGTGCCTCCGC 533
Db        434 GAGGAGAGGCGGGACCCGCGGGGCCCACCGGGCCTGCCGGGGAGTGCTCGGTGCCTCCGC 493
Qy        534 GATCCGCCTTCAGCGCCAAGCGCTCCGAGAGCCGGGTGCCTCCGCCGTCTGACGCACCCT 593
Db        494 GATCCGCCTTCAGCGCCAAGCGCTCCGAGAGCCGGGTGCCTCCGCCGTCTGACGCACCCT 553
Qy        594 TGCCCTTCGACCGCGTGCTGGTGAACGAGCAGGGACATTACGACGCCGTCACCGGCAAGT 653
Db        554 TGCCCTTCGACCGCGTGCTGGTGAACGAGCAGGGACATTACGACGCCGTCACCGGCAAGT 613
Qy        654 TCACCTGCCAGGTGCCTGGGGTCTACTACTTCGCCGTCCATGCCACCGTCTACCGGGCCA 713
Db        614 TCACCTGCCAGGTGCCTGGGGTCTACTACTTCGCCGTCCATGCCACCGTCTACCGGGCCA 673
Qy        714 GCCTGCAGTTTGATCTGGTGAAGAATGGCGAATCCATTGCCTCTTTCTTCCAGTTTTTCG 773
Db        674 GCCTGCAGTTTGATCTGGTGAAGAATGGCGAATCCATTGCCTCTTTCTTCCAGTTTTTCG 733
Qy        774 GGGGGTGGCCCAAGCCAGCCTCGCTCTCGGGGGGGGCCATGGTGAGGCTGGAGCCTGAGG 833
Db        734 GGGGGTGGCCCAAGCCAGCCTCGCTCTCGGGGGGGGCCATGGTGAGGCTGGAGCCTGAGG 793
```

```
Qy        834 ACCAAGTGTGGGTGCAGGTGGGTGTGGGTGACTACATTGGCATCTATGCCAGCATCAAGA 893
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        794 ACCAAGTGTGGGTGCAGGTGGGTGTGGGTGACTACATTGGCATCTATGCCAGCATCAAGA 853

Qy        894 CAGACAGCACCTTCTCCGGATTTCTGGTGTACTCCGACTGGCACAGCTCCCCAGTCTTTG 953
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        854 CAGACAGCACCTTCTCCGGATTTCTGGTGTACTCCGACTGGCACAGCTCCCCAGTCTTTG 913

Qy        954 CTTAGTGCCCACTGCAAAGTGAGCTCATGCTCTCACTCCTAGAAGGAGGGTGTGAGGCTG 1013
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        914 CTTAGTGCCCACTGCAAAGTGAGCTCATGCTCTCACTCCTAGAAGGAGGGTGTGAGGCTG 973

Qy       1014 ACAACCAGGTCATCCAGGAGGGCTGGCCCCCCCTGGAATATTGTGAATGACTAGGGAGGTG 1073
              |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        974 ACAACCTGGTCATCCAGGAGGGCTGGCCCCCCTGGAATATTGTGAATGACTAGGGAGGTG 1033

Qy       1074 GGGTAGAGCACTCTCCGTCCTGCTGCTGGCAAGGAATGGGAACAGTGGCTGTCTGCGATC 1133
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1034 GGGTAGAGCACTCTCCGTCCTGCTGCTGGCAAGGAATGGGAACAGTGGCTGTCTGCGATC 1093

Qy       1134 AGGTCTGGCAGCATGGGGCAGTGGCTGGATTTCTGCCCAAGACCAGAGGAGTGTGCTGTG 1193
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1094 AGGTCTGGCAGCATGGGGCAGTGGCTGGATTTCTGCCCAAGACCAGAGGAGTGTGCTGTG 1153

Qy       1194 CTGGCAAGTGTAAGTCCCCCAGTTGCTCTGGTCCAGGAGCCCACGGTGGGGTGCTCTCTT 1253
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1154 CTGGCAAGTGTAAGTCCCCCAGTTGCTCTGGTCCAGGAGCCCACGGTGGGGTGCTCTCTT 1213

Qy       1254 CCTGGTCCTCTGCTTCTCTGGATCCTCCCCACCCCCTCCTGCTCCTGGGGCCGGCCCTTT 1313
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1214 CCTGGTCCTCTGCTTCTCTGGATCCTCCCCACCCCCTCCTGCTCCTGGGGCCGGCCCTTT 1273

Qy       1314 TCTCAGAGATCACTCAATAAACCTAAGAACCCTCATAAAAAAAAAAAAAAAAAAA 1368
              ||||||||||||||||||||||||||||||||  ||||||||||||||||||||
Db       1274 TCTCAGAGATCACTCAATAAACCTAAGAACCCTCCAAAAAAAAAAAAAAAAAAAA 1328
```

RESULT 3
US-09-336-536-2
; Sequence 2, Application US/09336536
; Patent No. 6406884
; GENERAL INFORMATION:
; APPLICANT: Leiby, K.
; APPLICANT: McKay, C.
; APPLICANT: Bossone, S.
; TITLE OF INVENTION: SECRETED PROTEINS AND USES THEREOF
; FILE REFERENCE: 7853-144
; CURRENT APPLICATION NUMBER: US/09/336,536
; CURRENT FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 728
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-336-536-2

Query Match 52.1%; Score 717; DB 4; Length 728;
Best Local Similarity 99.9%; Pred. No. 1.4e-135;
Matches 728; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

```
Qy        227 ATGAGGCCACTCCTCGTCCTGCTGCTCCTGGGCCTGGCGGCCGGCTCGCCCCCACTGGAC 286
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGAGGCCACTCCTCGTCCTGCTGCTCCTGGGCCTGGCGGCCGGCTCGCCCCCACTGGAC 60

Qy        287 GACAACAAGATCCCCAGCCTCTGCCCGGGGCACCCCGGCCTTCCAGGCACGCCGGGCCAC 346
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GACAACAAGATCCCCAGCCTCTGCCCGGGGCACCCCGGCCTTCCAGGCACGCCGGGCCAC 120

Qy        347 CATGGCAGCCAGGGCTTGCCGGGCCGCGATGGCCGCGACGGCCGCGACGGCGCGCCCGGG 406
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CATGGCAGCCAGGGCTTGCCGGGCCGCGATGGCCGCGACGGCCGCGACGGCGCGCCCGGG 180

Qy        407 GCTCCGGGAGAGAAAGGCGAGGGCGGGAGGCCGGGACTGCCGGGACCTCGAGGGGACCCC 466
              |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        181 GCTCCGGGAGAGAAAGGCGAGGGCGGGAGG-CGGGACTGCCGGGACCTCGAGGGGACCCC 239

Qy        467 GGGCCGCGAGGAGAGGCGGGACCCGCGGGGCCCACCGGGCCTGCCGGGGAGTGCTCGGTG 526
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        240 GGGCCGCGAGGAGAGGCGGGACCCGCGGGGCCCACCGGGCCTGCCGGGGAGTGCTCGGTG 299

Qy        527 CCTCCGCGATCCGCCTTCAGCGCCAAGCGCTCCGAGAGCCGGGTGCCTCCGCCGTCTGAC 586
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        300 CCTCCGCGATCCGCCTTCAGCGCCAAGCGCTCCGAGAGCCGGGTGCCTCCGCCGTCTGAC 359

Qy        587 GCACCCTTGCCCTTCGACCGCGTGCTGGTGAACGAGCAGGGACATTACGACGCCGTCACC 646
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        360 GCACCCTTGCCCTTCGACCGCGTGCTGGTGAACGAGCAGGGACATTACGACGCCGTCACC 419

Qy        647 GGCAAGTTCACCTGCCAGGTGCCTGGGGTCTACTACTTCGCCGTCCATGCCACCGTCTAC 706
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        420 GGCAAGTTCACCTGCCAGGTGCCTGGGGTCTACTACTTCGCCGTCCATGCCACCGTCTAC 479

Qy        707 CGGGCCAGCCTGCAGTTTGATCTGGTGAAGAATGGCGAATCCATTGCCTCTTTCTTCCAG 766
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        480 CGGGCCAGCCTGCAGTTTGATCTGGTGAAGAATGGCGAATCCATTGCCTCTTTCTTCCAG 539

Qy        767 TTTTTCGGGGGGTGGCCCAAGCCAGCCTCGCTCTCGGGGGGGGCCATGGTGAGGCTGGAG 826
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        540 TTTTTCGGGGGGTGGCCCAAGCCAGCCTCGCTCTCGGGGGGGGCCATGGTGAGGCTGGAG 599

Qy        827 CCTGAGGACCAAGTGTGGGTGCAGGTGGGTGTGGGTGACTACATTGGCATCTATGCCAGC 886
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        600 CCTGAGGACCAAGTGTGGGTGCAGGTGGGTGTGGGTGACTACATTGGCATCTATGCCAGC 659

Qy        887 ATCAAGACAGACAGCACCTTCTCCGGATTTCTGGTGTACTCCGACTGGCACAGCTCCCCA 946
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        660 ATCAAGACAGACAGCACCTTCTCCGGATTTCTGGTGTACTCCGACTGGCACAGCTCCCCA 719

Qy        947 GTCTTTGCT 955
              |||||||||
Db        720 GTCTTTGCT 728
```

RESULT 4
US-09-336-536-8
; Sequence 8, Application US/09336536
; Patent No. 6406884
; GENERAL INFORMATION:
; APPLICANT: Leiby, K.
; APPLICANT: McKay, C.
; APPLICANT: Bossone, S.
; TITLE OF INVENTION: SECRETED PROTEINS AND USES THEREOF
; FILE REFERENCE: 7853-144
; CURRENT APPLICATION NUMBER: US/09/336,536
; CURRENT FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 1263
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-336-536-8

Query Match 50.3%; Score 692.8; DB 4; Length 1263;
Best Local Similarity 78.0%; Pred. No. 1.1e-130;
Matches 917; Conservative 0; Mismatches 237; Indels 22; Gaps 6;

```
Qy        189 GAGAGGACCCCGGCGTCCGGGCTCCCGGTGCCAGCGCTATGAGGCCACTCCTCGTCCTGC 248
              ||||||  || || ||| | ||  |||| | || | | |||||||||| || | ||| |
Db         97 GAGAGGTTCCTGGAGTCTGAGCCTCCGGGGTCACCACCATGAGGCCACTTCTTGCCCTTC 156

Qy        249 TGCTCCTGGGCCTGGCGGCCGGCTCGCCCCCACTGGACGACAACAAGATCCCCAGCCTCT 308
              |||| ||||| |||| | | ||||| || || ||||||||||||||||||||||||| |
Db        157 TGCTTCTGGGTCTGGTGTCAGGCTCTCCTCCTCTGGACGACAACAAGATCCCCAGCCTGT 216
```

```
Qy        309 GCCCGGGGCACCCCGGCCTTCCAGGCACGCCGGGCCACCATGGCAGCCAGGGCTTGCCGG 368
Db        217 GTCCCGGGCAGCCCGGCCTTCCAGGCACACCAGGTCACCATGGCAGCCAAGGCCTGCCTG 276

Qy        369 GCCGCGATGGCCGCGACGGCCGCGACGGCGCGCCCCGGGGCTCCGGGAGAGAAAGGCGAGG 428
Db        277 GCCGTGACGGCCGTGATGGCCGCGACGGTGCACCCGGAGCTCCGGGAGAGAAAGGCGAGG 336

Qy        429 GCGGGAGGCCGGGACTGCCGGGACCTCGAGGGGACCCCGGGCCGCGAGGAGAGGCGGGAC 488
Db        337 GCGGGAGACCGGGACTACCTGGCCCACGTGGGGAGCCCGGGCCGCGTGGAGAGGCAGGGC 396

Qy        489 CCGCGGGGCCCACCGGGCCTGCCGGGGAGTGCTCGGTGCCTCCGCGATCCGCCTTCAGCG 548
Db        397 CCATGGGGGCTATCGGGCCTGCGGGGGAGTGCTCGGTACCCCCACGATCAGCCTTCAGTG 456

Qy        549 CCAAGCGCTCCGAGAGCCGGGTGCCTCCGCCGTCTGACGCACCCTTGCCCTTCGACCGCG 608
Db        457 CCAAGCGATCCGAGAGCCGGGTACCTCCGCCAGCCGACACACCCCTACCTTTCGACCGTG 516

Qy        609 TGCTGGTGAACGAGCAGGGACATTACGACGCCGTCACCGGCAAGTTCACCTGCCAGGTGC 668
Db        517 TGCTGCTAAATGAGCAGGGCCATTACGACCCCACTACTGGCAAGTTCACCTGCCAAGTGC 576

Qy        669 CTGGGGTCTACTACTTCGCCGTCCATGCCACCGTCTACCGGGCCAGCCTGCAGTTTGATC 728
Db        577 CTGGCGTCTACTACTTTGCTGTGCACGCCACTGTCTACCGGGCCAGCTTGCAGTTTGATC 636

Qy        729 TGGTGAAGAATGGCGAATCCATTGCCTCTTTCTTCCAGTTTTTCGGGGGGTGGCCCAAGC 788
Db        637 TTGTCAAAAACGGGCAGTCCATCGCCTCTTTCTTCCAGTATTTTGGGGGGTGGCCCAAGC 696

Qy        789 CAGCCTCGCTCTCGGGGGGGGCCATGGTGAGGCTGGAGCCTGAGGACCAAGTGTGGGTGC 848
Db        697 CAGCCTCGCTCTCAGGGGGTGCGATGGTAAGGCTAGAACCTGAGGACCAGGTGTGGGTGC 756

Qy        849 AGGTGGGTGTGGGTGACTACATTGGCATCTATGCCAGCATCAAGACAGACAGCACCTTCT 908
Db        757 AGGTGGGCGTGGGTGATTACATTGGCATCTATGCCAGCATCAAGACAGACAGTACCTTCT 816

Qy        909 CCGGATTTCTGGTGTACTCCGACTGGCACAGCTCCCCAGTCTTTGCTTAGTGCCCACTGC 968
Db        817 CTGGATTTCTCGTCTATTCTGACTGGCACAGCTCCCCAGTCTTCGCTTAAAACACAGTGA 876

Qy        969 AAAGTGAGCTCATGCTCTCACTCCTAGAAGGAGGGTGTGAGGCTGACAACCAGGTCATCC 1028
Db        877 ACCCGGAGCTGGCACTTGCTCCTCAGTGGAGGGTGTGACACTAACCCGCGCAGCGCATAC 936

Qy       1029 AGGAGGGCTGGCCCCCCTGGAATATTGTGAATGACTAGGGAGGTGGGGTAGAGCACTCTC 1088
Db        937 CAGGAGGGCTGGCCCCCTGGAATATTGTGAATGACTTAGGAAGAGAGGGAG-CCACTTCC 995

Qy       1089 CGTCCTGCTGCTGGCAAGGAATGGGAACAGTGGCTGTCTGCGATCAGGTCTGGCAGCATG 1148
Db        996 AGTCCCACTGCTGGCAATGAATGGAGACA--GGCTGTCTGAGGTCA----AGACAGCGTG 1049

Qy       1149 GGGCAGTGGCTGGATTTCTGCCCAAGACCAGAGGAGTGTGCTGTGCTGGCAAGTGTAAGT 1208
Db       1050 GAGCAGTGGCTGGGTTTCTGCCCAGGACTTTA-GAATGCAGTAGGCTGGCAGCTGTGGGT 1108

Qy       1209 CCCCCAGTTGCTCTGGTCCAGGAGCCCACGGTGGGGTGCTCTCTTCCTGGTCCTCTG--- 1265
Db       1109 C-----------CTGGCCCAGGACTCCAAGGTGGGATGCTCCATTCCTAGTCCTGTGTCC 1157

Qy       1266 CTTCTCTGGATCCTCCCCACCCCCTCCTGCTCCTGGGGCCGGCCCTTTTCTCAGAGATCA 1325
Db       1158 CCTCTAGGTCCCTGACTCCATCTCTGCTGCTCCCAGGGCAGGCCTTTTTCTCAGAGGTCA 1217

Qy       1326 CTCAATAAACCTAAGAACCCTCATAAAAAAAAAAAA 1361
Db       1218 CTTAATAAACCTAAAATCCTCAAAAAAAAAAAAAA 1253
```

```
RESULT 5
US-09-188-930-218
; Sequence 218, Application US/09188930A
; Patent No. 6150502
; GENERAL INFORMATION:
;  APPLICANT: Watson, James D.
;  APPLICANT: Strachan, Lorna
;  APPLICANT: Sleeman, Matthew
;  APPLICANT: Onrust, Rene
;  APPLICANT: Murison, James Greg
;  TITLE OF INVENTION: Compositions Isolated From Skin Cells
;  TITLE OF INVENTION: and Methods For Their Use
;  FILE REFERENCE: 11000.1011c1
;  CURRENT APPLICATION NUMBER: US/09/188,930A
;  CURRENT FILING DATE: 1998-11-09
;  NUMBER OF SEQ ID NOS: 348
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 218
;   LENGTH: 1001
;   TYPE: DNA
;   ORGANISM: Rat
US-09-188-930-218

  Query Match           44.0%;  Score 606;  DB 3;  Length 1001;
  Best Local Similarity  78.6%;  Pred. No. 2.7e-113;
  Matches  739;  Conservative    0;  Mismatches  195;  Indels    6;  Gaps    1;

Qy         66 TGACGGCAGGGCAGGGGGCGCCTGGCCGGGGAGAAGCGCGGGGGCTGGAGCACCACCAAC 125
Db         31 TGTCAGCGAGGGCAGGGGCTGCCTGGTTGGGGTAGGAGTGGGAGCAGGGCCAGCAGGAGG 90

Qy        126 TGGAGGGTCCGGAGTAGCGAGCGCCCCGAAGGAGGCCATCGGGGAGCCGGGAGGGGGGAC 185
Db         91 GTCTGAGGAAGCCATTCAAAGCGAGCAGCTGGGAGAGCTGGGGAGCCGGGAAGGGCCTAC 150

Qy        186 TG------CGAGAGGACCCCGGCGTCCGGGCTCCCGGTGCCAGCGCTATGAGGCCACTCC 239
Db        151 AGACTACAAGAGAGGATCCTGGCGTCTGGGCCTCCTGGGTCATCACCATGAGGCCACTTC 210

Qy        240 TCGTCCTGCTGCTCCTGGGCCTGGCGGCCGGCTCGCCCCCACTGGACGACAACAAGATCC 299
Db        211 TTGCCCTGCTGCTTCTGGGTCTGGCATCAGGCTCTCCTCCTCTGGACGACAACAAGATCC 270

Qy        300 CCAGCCTCTGCCCGGGGCACCCCGGCCTTCCAGGCACGCCGGGCCACCATGGCAGCCAGG 359
Db        271 CCAGCCTGTGTCCCGGGCAGCCCGGCCTCCCAGGCACACCAGGCCACCACGGCAGCCAAG 330

Qy        360 GCTTGCCGGGCCGCGATGGCCGCGACGGCCGCGACGGCGCGCCCGGGGCTCCGGGAGAGA 419
Db        331 GCCTGCCTGGCCGTGACGGCCGTGATGGCCGCGACGGTGCACCCGGAGCTCCGGGAGAGA 390

Qy        420 AAGGCGAGGGCGGGAGGCCGGGACTGCCGGGACCTCGAGGGGACCCCGGGCCGCGAGGAG 479
Db        391 AAGGCGAGGGCGGGAGACCGGGACTACCTGGGCCACGTGGGGAGCCCGGGCCGCGTGGAG 450

Qy        480 AGGCGGGACCCGCGGGGCCCACCGGGCCTGCCGGGGAGTGCTCGGTGCCTCCGCGATCCG 539
Db        451 AGGCAGGACCTGTGGGGGCTATCGGGCCTGCGGGGGAGTGCTCGGTGCCCCCACGATCAG 510

Qy        540 CCTTCAGCGCCAAGCGCTCCGAGAGCCGGGTGCCTCCGCCGTCTGACGCACCCTTGCCCT 599
Db        511 CCTTCAGTGCCAAGCGATCAGAGAGCCGGGTACCTCCGCCAGCCGACACACCCCTACCCT 570

Qy        600 TCGACCGCGTGCTGGTGAACGAGCAGGGACATTACGACGCCGTCACCGGCAAGTTCACCT 659
Db        571 TCGACCGTGTGCTGCTCAATGAGCAGGGACATTACGATGCCACTACCGGCAAGTTCACCT 630

Qy        660 GCCAGGTGCCTGGGGTCTACTACTTCGCCGTCCATGCCACCGTCTACCGGGCCAGCCTGC 719
Db        631 GCCAAGTGCCTGGTGTCTACTACTTTGCTGTCCATGCCACTGTCTACCGGGCCAGCCTAC 690

Qy        720 AGTTTGATCTGGTGAAGAATGGCGAATCCATTGCCTCTTTCTTCCAGTTTTTCGGGGGGT 779
```

```
Db        691 AGTTTGATCTTGTCAAAAATGGCCAATCCATAGCTTCTTTCTTCCAGTTTTTTGGGGGGT 750

Qy        780 GGCCCAAGCCAGCCTCGCTCTCGGGGGGGGCCATGGTGAGGCTGGAGCCTGAGGACCAAG 839

Db        751 GGCCAAAGCCAGCCTCGCTCTCAGGGGGTGCGATGGTGAGGCTAGAACCTGAGGACCAGG 810

Qy        840 TGTGGGTGCAGGTGGGTGTGGGTGACTACATTGGCATCTATGCCAGCATCAAGACAGACA 899

Db        811 TATGGGTTCAGGTGGGTGTGGGTGATTACATTGGCATCTATGCCAGCATCAAAACAGACA 870

Qy        900 GCACCTTCTCCGGATTTCTGGTGTACTCCGACTGGCACAGCTCCCCAGTCTTTGCTTAGT 959

Db        871 GTACCTTCTCTGGATTTCTCGTCTATTCTGACTGGCACAGCTCCCCAGTCTTCGCTTAAA 930

Qy        960 GCCCACTGCAAAGTGAGCTCATGCTCTCACTCCTAGAAGG 999

Db        931 ATACAGTGAACCCGGAGCTGGCACTTGCTCCTAGTGGAGG 970
```

RESULT 6
US-09-188-930-30
; Sequence 30, Application US/09188930A
; Patent No. 6150502
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene
; APPLICANT: Murison, James Greg
; TITLE OF INVENTION: Compositions Isolated From Skin Cells
; TITLE OF INVENTION: and Methods For Their Use
; FILE REFERENCE: 11000.1011c1
; CURRENT APPLICATION NUMBER: US/09/188,930A
; CURRENT FILING DATE: 1998-11-09
; NUMBER OF SEQ ID NOS: 348
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 30
; LENGTH: 1015
; TYPE: DNA
; ORGANISM: Rat
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (545)...(545)
US-09-188-930-30

Query Match 43.9%; Score 605; DB 3; Length 1015;
Best Local Similarity 78.5%; Pred. No. 4.3e-113;
Matches 738; Conservative 0; Mismatches 196; Indels 6; Gaps 1;

```
Qy         66 TGACGGCAGGGCAGGGGGCGCCTGGCCGGGGAGAAGCGCGGGGGCTGGAGCACCACCAAC 125

Db         45 TGTCAGCGAGGGCAGGGGCTGCCTGGTTGGGGTAGGAGTGGGAGCAGGGCCAGCAGGAGG 104

Qy        126 TGGAGGGTCCGGAGTAGCGAGCGCCCCGAAGGAGGCCATCGGGGAGCCGGGAGGGGGGAC 185

Db        105 GTCTGAGGAAGCCATTCAAAGCGAGCAGCTGGGAGAGCTGGGGAGCCGGGAAGGGCCTAC 164

Qy        186 TG------CGAGAGGACCCCGGCGTCCGGGCTCCCGGTGCCAGCGCTATGAGGCCACTCC 239

Db        165 AGACTACAAGAGAGGATCCTGGCGTCTGGGCCTCCTGGGTCATCACCATGAGGCCACTTC 224

Qy        240 TCGTCCTGCTGCTCCTGGGCCTGGCGGCCGGCTCGCCCCCACTGGACGACAACAAGATCC 299

Db        225 TTGCCCTGCTGCTTCTGGGTCTGGCATCAGGCTCTCCTCCTCTGGACGACAACAAGATCC 284

Qy        300 CCAGCCTCTGCCCGGGGCACCCCGGCCTTCCAGGCACGCCGGGCCACCATGGCAGCCAGG 359

Db        285 CCAGCCTGTGTCCCGGGCAGCCCGGCCTCCCAGGCACACCAGGCCACCACGGCAGCCAAG 344

Qy        360 GCTTGCCGGGCCGCGATGGCCGCGACGGCCGCGACGGCGCGCCCGGGGCTCCGGGAGAGA 419

Db        345 GCCTGCCTGGCCGTGACGGCCGTGATGGCCGCGACGGTGCACCCGGAGCTCCGGGAGAGA 404

Qy        420 AAGGCGAGGGCGGGAGGCCGGGACTGCCGGGACCTCGAGGGGACCCCGGGCCGCGAGGAG 479

Db        405 AAGGCGAGGGCGGGAGACCGGGACTACCTGGGCCACGTGGGGAGCCCGGGCCGCGTGGAG 464

Qy        480 AGGCGGGACCCGCGGGGCCCACCGGGCCTGCCGGGGAGTGCTCGGTGCCTCCGCGATCCG 539

Db        465 AGGCAGGACCTGTGGGGGCTATCGGGCCTGCGGGGGAGTGCTCGGTGCCCCCACGATCAG 524

Qy        540 CCTTCAGCGCCAAGCGCTCCGAGAGCCGGGTGCCTCCGCCGTCTGACGCACCCTTGCCCT 599

Db        525 CCTTCAGTGCCAAGCGATCANAGAGCCGGGTACCTCCGCCAGCCGACACACCCCTACCCT 584

Qy        600 TCGACCGCGTGCTGGTGAACGAGCAGGGACATTACGACGCCGTCACCGGCAAGTTCACCT 659

Db        585 TCGACCGTGTGCTGCTCAATGAGCAGGGACATTACGATGCCACTACCGGCAAGTTCACCT 644

Qy        660 GCCAGGTGCCTGGGGTCTACTACTTCGCCGTCCATGCCACCGTCTACCGGGCCAGCCTGC 719

Db        645 GCCAAGTGCCTGGTGTCTACTACTTTGCTGTCCATGCCACTGTCTACCGGGCCAGCCTAC 704

Qy        720 AGTTTGATCTGGTGAAGAATGGCGAATCCATTGCCTCTTTCTTCCAGTTTTTCGGGGGGT 779

Db        705 AGTTTGATCTTGTCAAAAATGGCCAATCCATAGCTTCTTTCTTCCAGTTTTTTGGGGGGT 764

Qy        780 GGCCCAAGCCAGCCTCGCTCTCGGGGGGGGCCATGGTGAGGCTGGAGCCTGAGGACCAAG 839

Db        765 GGCCAAAGCCAGCCTCGCTCTCAGGGGGTGCGATGGTGAGGCTAGAACCTGAGGACCAGG 824

Qy        840 TGTGGGTGCAGGTGGGTGTGGGTGACTACATTGGCATCTATGCCAGCATCAAGACAGACA 899

Db        825 TATGGGTTCAGGTGGGTGTGGGTGATTACATTGGCATCTATGCCAGCATCAAAACAGACA 884

Qy        900 GCACCTTCTCCGGATTTCTGGTGTACTCCGACTGGCACAGCTCCCCAGTCTTTGCTTAGT 959

Db        885 GTACCTTCTCTGGATTTCTCGTCTATTCTGACTGGCACAGCTCCCCAGTCTTCGCTTAAA 944

Qy        960 GCCCACTGCAAAGTGAGCTCATGCTCTCACTCCTAGAAGG 999

Db        945 ATACAGTGAACCCGGAGCTGGCACTTGCTCCTAGTGGAGG 984
```

RESULT 7
US-09-336-536-9
; Sequence 9, Application US/09336536
; Patent No. 6406884
; GENERAL INFORMATION:
; APPLICANT: Leiby, K.
; APPLICANT: McKay, C.
; APPLICANT: Bossone, S.
; TITLE OF INVENTION: SECRETED PROTEINS AND USES THEREOF
; FILE REFERENCE: 7853-144
; CURRENT APPLICATION NUMBER: US/09/336,536
; CURRENT FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 729
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-336-536-9

Query Match 41.0%; Score 564.2; DB 4; Length 729;
Best Local Similarity 85.9%; Pred. No. 6.2e-105;
Matches 626; Conservative 0; Mismatches 103; Indels 0; Gaps 0;

```
Qy        227 ATGAGGCCACTCCTCGTCCTGCTGCTCCTGGGCCTGGCGGCCGGCTCGCCCCCACTGGAC 286

Db          1 ATGAGGCCACTTCTTGCCCTTCTGCTTCTGGGTCTGGTGTCAGGCTCTCCTCCTCTGGAC 60

Qy        287 GACAACAAGATCCCCAGCCTCTGCCCGGGGCACCCCGGCCTTCCAGGCACGCCGGGCCAC 346

Db         61 GACAACAAGATCCCCAGCCTGTGTCCCGGGCAGCCCGGCCTTCCAGGCACACCAGGTCAC 120
```

```
Qy        347 CATGGCAGCCAGGGCTTGCCGGGCCGCGATGGCCGCGACGGCCGCGACGGCGCGCCCCGGG 406
Db        121 CATGGCAGCCAAGGCCTGCCTGGCCGTGACGGCCGTGATGGCCGCGACGGTGCACCCGGA 180

Qy        407 GCTCCGGGAGAGAAAGGCGAGGGCGGGAGGCCGGGACTGCCGGGACCTCGAGGGGACCCC 466
Db        181 GCTCCGGGAGAGAAAGGCGAGGGCGGGAGACCGGGACTACCTGGCCCACGTGGGGAGCCC 240

Qy        467 GGGCCGCGAGGAGAGGCGGGACCCGCGGGGCCCACCGGGCCTGCCGGGGAGTGCTCGGTG 526
Db        241 GGGCCGCGTGGAGAGGCAGGGCCCATGGGGGCTATCGGGCCTGCGGGGGAGTGCTCGGTA 300

Qy        527 CCTCCGCGATCCGCCTTCAGCGCCAAGCGCTCCGAGAGCCGGGTGCCTCCGCCGTCTGAC 586
Db        301 CCCCCACGATCAGCCTTCAGTGCCAAGCGATCCGAGAGCCGGGTACCTCCGCCAGCCGAC 360

Qy        587 GCACCCTTGCCCTTCGACCGCGTGCTGGTGAACGAGCAGGGACATTACGACGCCGTCACC 646
Db        361 ACACCCCTACCTTTCGACCGTGTGCTGCTAAATGAGCAGGGCCATTACGACCCCACTACT 420

Qy        647 GGCAAGTTCACCTGCCAGGTGCCTGGGGTCTACTACTTCGCCGTCCATGCCACCGTCTAC 706
Db        421 GGCAAGTTCACCTGCCAAGTGCCTGGCGTCTACTACTTTGCTGTGCACGCCACTGTCTAC 480

Qy        707 CGGGCCAGCCTGCAGTTTGATCTGGTGAAGAATGGCGAATCCATTGCCTCTTTCTTCCAG 766
Db        481 CGGGCCAGCTTGCAGTTTGATCTTGTCAAAAACGGGCAGTCCATCGCCTCTTTCTTCCAG 540

Qy        767 TTTTTCGGGGGGTGGCCCAAGCCAGCCTCGCTCTCGGGGGGGGCCATGGTGAGGCTGGAG 826
Db        541 TATTTTGGGGGGTGGCCCAAGCCAGCCTCGCTCTCAGGGGGTGCGATGGTAAGGCTAGAA 600

Qy        827 CCTGAGGACCAAGTGTGGGTGCAGGTGGGTGTGGGTGACTACATTGGCATCTATGCCAGC 886
Db        601 CCTGAGGACCAGGTGTGGGTGCAGGTGGGCGTGGGTGATTACATTGGCATCTATGCCAGC 660

Qy        887 ATCAAGACAGACAGCACCTTCTCCGGATTTCTGGTGTACTCCGACTGGCACAGCTCCCCA 946
Db        661 ATCAAGACAGACAGTACCTTCTCTGGATTTCTCGTCTATTCTGACTGGCACAGCTCCCCA 720

Qy        947 GTCTTTGCT 955
Db        721 GTCTTCGCT 729


RESULT 8
US-09-140-804-10
; Sequence 10, Application US/09140804
; Patent No. 6197930
; GENERAL INFORMATION:
;  APPLICANT: Sheppard, Paul O.
;  APPLICANT: Humes, Jacqueline M.
;  TITLE OF INVENTION: ADIPOCYTE-SPECIFIC PROTEIN HOMOLOGS
;  FILE REFERENCE: 97-49
;  CURRENT APPLICATION NUMBER: US/09/140,804
;  CURRENT FILING DATE: 1998-08-26
;  EARLIER APPLICATION NUMBER: 60/056,983
;  EARLIER FILING DATE: 1997-08-26
;  NUMBER OF SEQ ID NOS: 47
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 10
;  LENGTH: 729
;  TYPE: DNA
;  ORGANISM: Artificial Sequence
;  FEATURE:
;  OTHER INFORMATION: Degenerate nucleotide sequence encoding the zsig39
;  OTHER INFORMATION: polypeptide of SEQ ID NO:2.
US-09-140-804-10

  Query Match             36.6%;  Score 504;  DB 4;  Length 729;
  Best Local Similarity   57.6%;  Pred. No. 7.5e-93;
  Matches  419;  Conservative  145;  Mismatches  164;  Indels    0;  Gaps    0;

Qy        227 ATGAGGCCACTCCTCGTCCTGCTGCTCCTGGGCCTGGCGGCCGGCTCGCCCCCACTGGAC 286
Db          1 ATGMGNCCNYTNYTNGTNYTNYTNYTNYTNGGNYTNGCNGCNGGNWSNCCNCCNYTNGAY 60

Qy        287 GACAACAAGATCCCCAGCCTCTGCCCGGGGCACCCCGGCCTTCCAGGCACGCCGGGCCAC 346
Db         61 GAYAAYAARATHCCNWSNYTNTGYCCNGGNCAYCCNGGNYTNCCNGGNACNCCNGGNCAY 120

Qy        347 CATGGCAGCCAGGGCTTGCCGGGCCGCGATGGCCGCGACGGCCGCGACGGCGCGCCCGGG 406
Db        121 CAYGGNWSNCARGGNYTNCCNGGNMGNGAYGGNMGNGAYGGNMGNGAYGGNGCNCCNGGN 180

Qy        407 GCTCCGGGAGAGAAAGGCGAGGGCGGGAGGCCGGGACTGCCGGGACCTCGAGGGGACCCC 466
Db        181 GCNCCNGGNGARAARGGNGARGGNGGNMGNCCNGGNYTNCCNGGNCCNMGNGGNGAYCCN 240

Qy        467 GGGCCGCGAGGAGAGGCGGGACCCGCGGGGCCCACCGGGCCTGCCGGGGAGTGCTCGGTG 526
Db        241 GGNCCNMGNGGNGARGCNGGNCCNGCNGGNCCNACNGGNCCNGCNGGNGARTGYWSNGTN 300

Qy        527 CCTCCGCGATCCGCCTTCAGCGCCAAGCGCTCCGAGAGCCGGGTGCCTCCGCCGTCTGAC 586
Db        301 CCNCCNMGNWSNGCNTTYWSNGCNAARMGNWSNGARWSNMGNGTNCCNCCNCCNWSNGAY 360

Qy        587 GCACCCTTGCCCTTCGACCGCGTGCTGGTGAACGAGCAGGGACATTACGACGCCGTCACC 646
Db        361 GCNCCNYTNCCNTTYGAYMGNGTNYTNGTNAAYGARCARGGNCAYTAYGAYGCNGTNACN 420

Qy        647 GGCAAGTTCACCTGCCAGGTGCCTGGGGTCTACTACTTCGCCGTCCATGCCACCGTCTAC 706
Db        421 GGNAARTTYACNTGYCARGTNCCNGGNGTNTAYTAYTTYGCNGTNCAYGCNACNGTNTAY 480

Qy        707 CGGGCCAGCCTGCAGTTTGATCTGGTGAAGAATGGCGAATCCATTGCCTCTTTCTTCCAG 766
Db        481 MGNGCNWSNYTNCARTTYGAYYTNGTNAARAAYGGNGARWSNATHGCNWSNTTYTTYCAR 540

Qy        767 TTTTTCGGGGGGTGGCCCAAGCCAGCCTCGCTCTCGGGGGGGGCCATGGTGAGGCTGGAG 826
Db        541 TTYTTYGGNGGNTGGCCNAARCCNGCNWSNYTNWSNGGNGGNGCNATGGTNMGNYTNGAR 600

Qy        827 CCTGAGGACCAAGTGTGGGTGCAGGTGGGTGTGGGTGACTACATTGGCATCTATGCCAGC 886
Db        601 CCNGARGAYCARGTNTGGGTNCARGTNGGNGTNGGNGAYTAYATHGGNATHTAYGCNWSN 660

Qy        887 ATCAAGACAGACAGCACCTTCTCCGGATTTCTGGTGTACTCCGACTGGCACAGCTCCCCA 946
Db        661 ATHAARACNGAYWSNACNTTYWSNGGNTTYYTNGTNTAYWSNGAYTGGCAYWSNWSNCCN 720

Qy        947 GTCTTTGC 954
Db        721 GTNTTYGC 728


RESULT 9
US-09-336-536-74
; Sequence 74, Application US/09336536
; Patent No. 6406884
; GENERAL INFORMATION:
;  APPLICANT: Leiby, K.
;  APPLICANT: McKay, C.
;  APPLICANT: Bossone, S.
;  TITLE OF INVENTION: SECRETED PROTEINS AND USES THEREOF
;  FILE REFERENCE: 7853-144
;  CURRENT APPLICATION NUMBER: US/09/336,536
;  CURRENT FILING DATE: 1999-06-18
;  NUMBER OF SEQ ID NOS: 75
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 74
;  LENGTH: 601
;  TYPE: DNA
;  ORGANISM: Rattus norvegicus
;  FEATURE:
```

```
;   NAME/KEY: modified_base
;   LOCATION: all "n" positions
;   OTHER INFORMATION: n=a, c, g, or t
US-09-336-536-74

  Query Match             18.7%;  Score 257.8;  DB 4;  Length 601;
  Best Local Similarity   72.0%;  Pred. No. 1.8e-43;
  Matches  365;  Conservative    0;  Mismatches  134;  Indels    8;  Gaps    2;

Qy         66 TGACGGCAGGGCAGGGGGCGCCTGGCCGGGGAGAAGCGCGGGGGCTGGAGCACCACCAAC 125
              || | ||  ||   |||||  |  |   |||  | | | ||| || ||  || ||  |
Db         96 TGTCAGCGAGGGCAGGGGCTGCCTGGTTGGGGTAGGAGTGGGAGCAGGGCCAGCAGGAGG 155

Qy        126 TGGAGGGTCCGGAGTAGCGAGCGCCCCGAAGGAGGCCATCGGGGAGCCGGGAGGGGGGAC 185
              | |   |   |    ||||  | |  ||  |   | |||   |_|| ||||   ||
Db        156 GTCTGAGGAAGCCATTCAAAGCGAGCAGCTGGGAGAGCTGGGGAGCCGGGAAGGGCCTAC 215

Qy        186 TG------CGAGAGGACCCCGGCGTCCGGGCTCCCGGTGCCAGCGCTATGAGGCCACTCC 239
               |       ||||||| || |||||| ||||  || | | || | | ||||||||||| |
Db        216 AGACTACAAGAGAGGATCCTGGCGTCTGGGCCTCCTGGGTCATCACCATGAGGCCACTTC 275

Qy        240 TCGTCCTGCTGCTCCTGGGCCTGGCGGCCGGCTCGCCCCCACTGGACGACAACAAGATCC 299
              | | |||||||||| ||||| |||||  | ||||| || || ||||||||||||||||||
Db        276 TTGCCCTGCTGCTTCTGGGTCTGGCATCAGGCTCTCCTCCTCTGGACGACAACAAGATCC 335

Qy        300 CCAGCCTCTGCCCGGGGCACCCCGGCCTTCCAGGCACGCCGGGCCACCATGGCAGCCAGG 359
              |||||||  || ||  ||||| |||||||| |||||||| || |||||||| ||||||| |
Db        336 CCAGCCTGTGTCCCGGGCAGCCCGGCCTCCCAGGCACACCAGGCCACCACGGCAGCCAAG 395

Qy        360 GCTTGCCGGGCCGCGATGGCCGCGACGGCCGCGACGGCGCGCCCGGGGCTCCGGGAGAGA 419
              || |||| ||||| || ||||| || |||||||||| || ||||| ||||||||||||
Db        396 GCCTGCCTGGCCGTGACGGCCGTGATGGCCGCGACGGTGCACCCGGAGCTCCGGGAGAGA 455

Qy        420 AAGGCGAGGGCGGGAGGCCGGGACTGCCGGGACCTCGAGGGGACCCCGGGCCGCGAGGAG 479
              ||||||||||||||| |||||||| || || || ||  |||| ||||||||||| ||||
Db        456 AAGGCGAGGGCGGGAGACCGGGACTACCTGGGCCACGTNGGGAGCCCGGGCCGCCGTGGAG 515

Qy        480 AGGCGGGACCCGCGGGGCCCACCGGGCCTGCCGGGGAGTGCTCGGTGCCTCCGCGATCCG 539
              |||| ||||| | |||| | | ||||||||  ||||| ||||||||||  ||      |
Db        516 AGGCAGGACCTGTGGGGGCTATCGGGCCTGGNGGGGAATGCTCGGTGC--CCCACGATCA 573

Qy        540 CCTTCAGCGCCAAGCGCTCCGAGAGCC 566
              |||||| |||||||| || || ||||
Db        574 GCTTCAGTGCCAAGCGATCAGAAAGCC 600


RESULT 10
US-09-188-930-26
; Sequence 26, Application US/09188930A
; Patent No. 6150502
; GENERAL INFORMATION:
;  APPLICANT: Watson, James D.
;  APPLICANT: Strachan, Lorna
;  APPLICANT: Sleeman, Matthew
;  APPLICANT: Onrust, Rene
;  APPLICANT: Murison, James Greg
;  TITLE OF INVENTION: Compositions Isolated From Skin Cells
;  TITLE OF INVENTION: and Methods For Their Use
;  FILE REFERENCE: 11000.1011c1
;  CURRENT APPLICATION NUMBER: US/09/188,930A
;  CURRENT FILING DATE: 1998-11-09
;  NUMBER OF SEQ ID NOS: 348
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 26
;   LENGTH: 393
;   TYPE: DNA
;   ORGANISM: Rat
US-09-188-930-26

  Query Match             10.9%;  Score 149.6;  DB 3;  Length 393;
  Best Local Similarity   69.1%;  Pred. No. 8.6e-22;
  Matches  250;  Conservative    0;  Mismatches  104;  Indels    8;  Gaps    3;

Qy         66 TGACGGCAGGGCAGGGGGCGCCTGGCCGGGGAGAAGCGCGGGGGCTGGAGCACCACCAAC 125
              || | ||  ||   |||||  |  |   |||  | | | ||| || ||  || ||  |
Db         33 TGTCAGCGAGGGCAGGGGCTGCCTGGTTGGGGTAGGAGTGGGAGCAGGGCCAGCAGGAGG 92

Qy        126 TGGAGGGTCCGGAGTAGCGAGCGCCCCGAAGGAGGCCATCGGGGAGCCGGGAGGGGGGAC 185
              | |   |   |    ||||  | |  ||  |   | |||   | || ||||   ||
Db         93 GTCTGAGGAAGCCATTCAAAGCGAGCAGCTGGGAGAGCTGGGGAGCCGGGAAGGGCCTAC 152

Qy        186 TG------CGAGAGGACCCCGGCGTCCGGGCTCCCGGTGCCAGCGCTATGAGGCCACTCC 239
               |       ||||||| || |||||| ||||  || | | || | | ||||||||||| |
Db        153 AGACTACAAGAGAGGATCCTGGCGTCTGGGCCTCCTGGGTCATCACCATGAGGCCACTTC 212

Qy        240 TCGTCCTGCTGCTCCTGGGCCTGGCGGCCGGCTCGCCCCCACTGGACGACAACAAGATCC 299
              | | |||||||||| ||||| |||||  | ||||| || || ||||||||||||||||||
Db        213 TTGCCCTGCTGCTTCTGGGTCTGGCATCAGGCTCTCCTCCTCTGGACGACAACAAGATCC 272

Qy        300 CCAGCCTCTGCCCGGGGCACCCCGGCCTTCCAGGCACGCCGGGCCACCATGGCAGCCAGG 359
              |||||||  || ||  ||||| |||||||| |||||||| || |||||||| ||||||| |
Db        273 CCAGCCTGTGTCCCGGGCAGCCCGGCCTCCCAGGCACACCAGGCCACCACGGCAGCCAAG 332

Qy        360 GCTTGCCGGGCCGCGATGGCCGCGACGGCCGCGACGGCGCGCCCGGGGCTCCGGGAGAGA 419
              || |||| ||||| || ||||  || |||||||||| || ||||| | ||||||||||
Db        333 GCCTGCCTGGCCGTGACGGCC-TGATGGCCGCGACGGTGCACCCGGAG-TCCGGGAGAGA 390

Qy        420 AA 421
              ||
Db        391 AA 392


RESULT 11
US-08-463-911-1
; Sequence 1, Application US/08463911
; Patent No. 5869330
;  GENERAL INFORMATION:
;    APPLICANT:  Scherer, Philipp E.
;    APPLICANT:  Lodish, Harvey F.
;    TITLE OF INVENTION:  A NOVEL SERUM PROTEIN PRODUCED
;    TITLE OF INVENTION:  EXCLUSIVELY IN ADIPOCYTES
;    NUMBER OF SEQUENCES:  7
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Hamilton, Brook, Smith & Reynolds, P.C.
;      STREET:  Two Militia Drive
;      CITY:  Lexington
;      STATE:  Massachusetts
;      COUNTRY:  USA
;      ZIP:  02173
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/463,911
;      FILING DATE:
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Granahan, Patricia
;      REGISTRATION NUMBER:  32,227
;      REFERENCE/DOCKET NUMBER:  WHI95-05
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 861-6240
;      TELEFAX:  (617) 861-9540
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1276 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  46..786
US-08-463-911-1
```

```
  Query Match             6.6%;  Score 90.6;  DB 2;  Length 1276;
  Best Local Similarity   51.5%;  Pred. No. 7.3e-10;
  Matches  316;  Conservative    0;  Mismatches  279;  Indels   18;  Gaps    4;

Qy        333 GCACGCCGGGCCACCATGGCAGCCAGGGCTTGCCGGGCCGCGATGGCCGCGACGGCCGCG 392
              | | | | ||| || || |  ||| |||      ||||| || ||| | |
Db        170 GGATGGCAGGCATCCCAGGACATCCTGGCCACAATGGCACACCAGGCCGTGATGGCAGAG 229

Qy        393 ACGGCGCGCCCGGGGCTCCGGGAGAGAAAGGCGAGGGCGGGAGGCCGGGACTGCCGGGAC 452
              | ||| | || || |    ||||||||||| || |  ||    | || |    |||
Db        230 ATGGCACTCCTGGAGAGAAGGGAGAGAAAGGAGATGCAGGTCTTCTTGGTCCTAAGGGTG 289

Qy        453 CTCGAGGGGACCCCGGGCCGCGAGGAGAGGCGGGACCCGCGGGGCCCACCGGGCCTGCCG 512
              . ||| ||    ||   | ||||  |  || ||   |||   | ||||  |  | |
Db        290 AGACAGGAGATGTTGGAATGACAGGAGCTGAAGGGCCACGGGGCTTCCCCGGAACCCCTG 349

Qy        513 GGGAGTGCTCGGTGCCTCCGCGATCCGCCTTCAG----CGCCAAGCGCTCCGAGAGCCGG 568
              |   |       | ||||     | |||| |  |      |||  |||| || | | |
Db        350 GCAGGAAAGGAGAGCCTGGAGAAGCCGCTTATATGTATCGCTCAGCGTTCAGTGTGGGGC 409

Qy        569 GTGCCTCCGCCGTCTG--------ACGCACCCTTGCCCTTCGACCGCGTGCTGGTGAACG 620
              |   ||  ||||          | | |||| | | |||      |  |    |||
Db        410 TGGAGACCCGCGTCACTGTTCCCAATGTACCCATTCGCTTTACTAAGATCTTCTACAACC 469

Qy        621 AGCAGGGACATTACGACGCCGTCACCGGCAAGTTCACCTGCCAGGTGCCTGGGGTCTACT 680
              | |||   ||||| |||| |  ||| |||||||||  |||| |  | || ||  ||||||
Db        470 AACAGAATCATTATGACGGCAGCACTGGCAAGTTCTACTGCAACATTCCGGGACTCTACT 529

Qy        681 ACTTCGCCGTCCATGCCACCGTCTACCGGGCCAGCCTGCAGTTTGATCTGGTGAAGAATG 740
              ||||| |   |||  ||| || ||| |       || || |    ||  | |||| |
Db        530 ACTTCTCTTACCACATCACGGTGTACATGAAAGATGTGAAGGTGAGCCTCTTCAAGAAGG 589

Qy        741 GCGAATCCATTGCCTCTTTCTTCCAGTTTTTCGGGGGGTGGCCCAAGCCAGCCTCGCTCT 800
              | |  || ||  ||    || | |    |    ||    |  ||    | |  |  ||
Db        590 ACAAGGCCGTTCTCTTCACCTACGACCAGTATCAGGAAAAG---AATGTGGACCAGGCCT 646

Qy        801 CGGGGGGGGCCATGGTGAGGCTGGAGCCTGAGGACCAAGTGTGGGTGCAGGTG---GGTG 857
              | ||    |   | |   ||||||   |  |||||||| ||| | ||||||   || |
Db        647 CTGGCTCTGTGCTCCTCCATCTGGAGGTGGGAGACCAAGTCTGGCTCCAGGTGTATGGGG 706

Qy        858 TGGGTGACTACATTGGCATCTATGCCAGCATCAAGACAGACAGCACCTTCTCCGGATTTC 917
              || ||| ||| |||  |||||||    | |   |  |||   || ||  | || ||||
Db        707 ATGGGGACCACAATGGACTCTATGCAGATAACGTCAACGACTCTACATTTACTGGCTTTC 766

Qy        918 TGGTGTACTCCGA 930
              |  | |||   ||
Db        767 TTCTCTACCATGA 779


RESULT 12
US-08-463-911-6
; Sequence 6, Application US/08463911
; Patent No. 5869330
;  GENERAL INFORMATION:
;    APPLICANT:  Scherer, Philipp E.
;    APPLICANT:  Lodish, Harvey F.
;    TITLE OF INVENTION:  A NOVEL SERUM PROTEIN PRODUCED
;    TITLE OF INVENTION:  EXCLUSIVELY IN ADIPOCYTES
;    NUMBER OF SEQUENCES:  7
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Hamilton, Brook, Smith & Reynolds, P.C.
;      STREET:  Two Militia Drive
;      CITY:  Lexington
;      STATE:  Massachusetts
;      COUNTRY:  USA
;      ZIP:  02173
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/463,911
;      FILING DATE:
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Granahan, Patricia
;      REGISTRATION NUMBER:  32,227
;      REFERENCE/DOCKET NUMBER:  WHI95-05
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 861-6240
;      TELEFAX:  (617) 861-9540
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1313 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  73..804
US-08-463-911-6

  Query Match             6.0%;  Score 82.2;  DB 2;  Length 1313;
  Best Local Similarity   50.1%;  Pred. No. 3.6e-08;
  Matches  316;  Conservative    0;  Mismatches  303;  Indels   12;  Gaps    4;

Qy        321 CCGGCCTTCCAGGCACGCCGGGCCACCATGGCAGCCAGGGCTTGCCGGGCCGCGATGGCC 380
              | ||| | |||||   ||||||||  ||||   ||  |||       ||| | ||||||
Db        194 CGGGCATCCCAGGGCATCCGGGCCATAATGGGGCCCCAGGCCGTGATGGCAGAGATGGCA 253

Qy        381 GCGACGGCCGCGACGGCGCGCCCCGGGGCTCCGGGAGAGAAAGGCGAGGGCGGGAGGCCGG 440
              |   ||      | || | |   || | ||| ||    |  ||          ||
Db        254 CCCCTGGTGAGAAGGGTGAGAAAGGAGATCCAGGTCTTATTGGTCCTAAGGGAGACATCG 313

Qy        441 GACTGCCGGGACCTCGAGGGGACCCCGGGCCGCGAGGAGAGGCGGGACCCGCGGGGCCCA 500
              |     | |||   | ||||     || || |||||     |||||  |   ||    |
Db        314 GTGAAACCGGAGTACCCGGGGCTGAAGGTCCCCGAGGCTTTCCGGGAATCCAAGGCAGGA 373

Qy        501 CCGGGCCTGCCGGGGAGTGCTCGGTGCCTCCGCGATCCGCCTTCAGCGCCAAGCGCTCCG 560
              ||      | || ||  |  |              || || || ||||| |   | | |  |
Db        374 AAGGAGAACCTGGAGAAGGTGCCTATGTATACCGCTCAGCATTCAGTGT---GGGATTGG 430

Qy        561 AGAGCCGGGTGCCTCCGCCGTCTGACGCACCCTTGCCCTTCGACCGCGTGCTGGTGAACG 620
              |||     ||  ||   |  |  ||   ||| | | |||    |   |  |    ||
Db        431 AGACTTACGTTACTATCC---CCAACATGCCCATTCGCTTTACCAAGATCTTCTACAATC 487

Qy        621 AGCAGGGACATTACGACGCCGTCACCGGCAAGTTCACCTGCCAGGTGCCTGGGGTCTACT 680
              ||||    || || || | |  ||| || || |||  |||| |  | |||||| | ||||
Db        488 AGCAAAACCACTATGATGGCTCCACTGGTAAATTCCACTGCAACATTCCTGGGCTGTACT 547

Qy        681 ACTTCGCCGTCCATGCCACCGTCTACCGGGCCAGCCTGCAGTTTGATCTGGTGAAGAATG 740
              |||| |||  |||   ||| |||||  |       || || |    ||  | ||||| |
Db        548 ACTTTGCCTACCACATCACAGTCTATATGAAGGATGTGAAGGTCAGCCTCTTCAAGAAGG 607

Qy        741 GCGAATCCATTGCCTCTTTCTTCCAGTTTTTCGGGGGGTGGCCCAAGCCAGCCTCGCTCT 800
              | |  | ||  ||    ||  |  | ||         ||    | |  ||  ||
Db        608 ACAAGGCTATGCTCTTCACCTATGATCAGTACCAGGAAAAT---AATGTGGACCAGGCCT 664

Qy        801 CGGGGGGGGCCATGGTGAGGCTGGAGCCTGAGGACCAAGTGTGGGTGCAGGTGGGTGTG- 859
              | ||    |  | ||  ||||||   |  |||||||| ||| | ||||||  || |
Db        665 CCGGCTCTGTGCTCCTGCATCTGGAGGTGGGCGACCAAGTCTGGCTCCAGGTGTATGGGG 724

Qy        860 --GGTGACTACATTGGCATCTATGCCAGCATCAAGACAGACAGCACCTTCTCCGGATTTC 917
                || ||    | |||  |||||||     |  | |  |||  ||||||| | || ||||
Db        725 AAGGAGAGCGTAATGGACTCTATGCTGATAATGACAATGACTCCACCTTCACAGGCTTTC 784

Qy        918 TGGTGTACTCCGACTGGCACAGCTCCCCAGT 948
              |  | |||   |||     || | || ||| |
Db        785 TTCTCTACCATGACACCAACTGATCACCACT 815


RESULT 13
US-09-140-804-9
```

```
; Sequence 9, Application US/09140804
; Patent No. 6197930
; GENERAL INFORMATION:
;  APPLICANT: Sheppard, Paul O.
;  APPLICANT: Humes, Jacqueline M.
;  TITLE OF INVENTION: ADIPOCYTE-SPECIFIC PROTEIN HOMOLOGS
;  FILE REFERENCE: 97-49
;  CURRENT APPLICATION NUMBER: US/09/140,804
;  CURRENT FILING DATE: 1998-08-26
;  EARLIER APPLICATION NUMBER: 60/056,983
;  EARLIER FILING DATE: 1997-08-26
;  NUMBER OF SEQ ID NOS: 47
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 9
;   LENGTH: 4517
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-140-804-9

  Query Match             6.0%;  Score 82.2;  DB 4;  Length 4517;
  Best Local Similarity   50.1%;  Pred. No. 4.4e-08;
  Matches  316;  Conservative    0;  Mismatches  303;  Indels   12;  Gaps    4;

Qy        321 CCGGCCTTCCAGGCACGCCGGGCCACCATGGCAGCCAGGGCTTGCCGGGCCGCGATGGCC 380
Db        148 CGGGCATCCCAGGGCATCCGGGCCATAATGGGGCCCCAGGCCGTGATGGCAGAGATGGCA 207

Qy        381 GCGACGGCCGCGACGGCGCGCCCGGGGCTCCGGGAGAGAAAGGCGAGGGCGGGAGGCCGG 440
Db        208 CCCCTGGTGAGAAGGGTGAGAAAGGAGATCCAGGTCTTATTGGTCCTAAGGGAGACATCG 267

Qy        441 GACTGCCGGGACCTCGAGGGGACCCCGGGCCGCGAGGAGAGGCGGGACCCGCGGGGCCCA 500
Db        268 GTGAAACCGGAGTACCCGGGGCTGAAGGTCCCCGAGGCTTTCCGGGAATCCAAGGCAGGA 327

Qy        501 CCGGGCCTGCCGGGGAGTGCTCGGTGCCTCCGCGATCCGCCTTCAGCGCCAAGCGCTCCG 560
Db        328 AAGGAGAACCTGGAGAAGGTGCCTATGTATACCGCTCAGCATTCAGTGT---GGGATTGG 384

Qy        561 AGAGCCGGGTGCCTCCGCCGTCTGACGCACCCTTGCCCTTCGACCGCGTGCTGGTGAACG 620
Db        385 AGACTTACGTTACTATCC---CCAACATGCCCATTCGCTTTACCAAGATCTTCTACAATC 441

Qy        621 AGCAGGGACATTACGACGCCGTCACCGGCAAGTTCACCTGCCAGGTGCCTGGGGTCTACT 680
Db        442 AGCAAAACCACTATGATGGCTCCACTGGTAAATTCCACTGCAACATTCCTGGGCTGTACT 501

Qy        681 ACTTCGCCGTCCATGCCACCGTCTACCGGGCCAGCCTGCAGTTTGATCTGGTGAAGAATG 740
Db        502 ACTTTGCCTACCACATCACAGTCTATATGAAGGATGTGAAGGTCAGCCTCTTCAAGAAGG 561

Qy        741 GCGAATCCATTGCCTCTTTCTTCCAGTTTTTCGGGGGGTGGCCCAAGCCAGCCTCGCTCT 800
Db        562 ACAAGGCTATGCTCTTCACCTATGATCAGTACCAGGAAAAT---AATGTGGACCAGGCCT 618

Qy        801 CGGGGGGGGCCATGGTGAGGCTGGAGCCTGAGGACCAAGTGTGGGTGCAGGTGGGTGTG- 859
Db        619 CCGGCTCTGTGCTCCTGCATCTGGAGGTGGGCGACCAAGTCTGGCTCCAGGTGTATGGGG 678

Qy        860 --GGTGACTACATTGGCATCTATGCCAGCATCAAGACAGACAGCACCTTCTCCGGATTTC 917
Db        679 AAGGAGAGCGTAATGGACTCTATGCTGATAATGACAATGACTCCACCTTCACAGGCTTTC 738

Qy        918 TGGTGTACTCCGACTGGCACAGCTCCCCAGT 948
Db        739 TTCTCTACCATGACACCAACTGATCACCACT 769


RESULT 14
US-09-118-408-23
; Sequence 23, Application US/09118408A
; Patent No. 6265544
; GENERAL INFORMATION:
;  APPLICANT: Sheppard, Paul O.
;  TITLE OF INVENTION: ADIPOCYTE-SPECIFIC PROTEIN HOMOLOGS
;  FILE REFERENCE: 97-30
;  CURRENT APPLICATION NUMBER: US/09/118,408A
;  CURRENT FILING DATE: 1998-07-17
;  EARLIER APPLICATION NUMBER: 60/053,154
;  EARLIER FILING DATE: 1997-07-18
;  NUMBER OF SEQ ID NOS: 47
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 23
;   LENGTH: 843
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Degenerate nucleotide sequence encoding zsig37
;   OTHER INFORMATION: polypeptide
;   FEATURE:
;   NAME/KEY: variation
;   LOCATION: (1)...(843)
;   OTHER INFORMATION: Each N is independently any nucleotide.
US-09-118-408-23

  Query Match             5.6%;  Score 77.4;  DB 4;  Length 843;
  Best Local Similarity   30.0%;  Pred. No. 3e-07;
  Matches  209;  Conservative   99;  Mismatches  384;  Indels    5;  Gaps    2;

Qy        233 CCACTCCTCGTCCTGCTGCTCCTGGGCCTGGCGGCCGGCTCGCCCCCACTGGACGACAAC 292
Db        131 CNWSNCCNCCNGAYCAYGCNGARMGNGCNGARGARCARCAYGARAARTAYMGNCCNWSNC 190

Qy        293 AAGATCCCCAGCCTCTGCCCGGGGCACCCCGGCCTTCCAGGCACGCCGGGCCACCATGGC 352
Db        191 ARGAYCARGGNYTNCCNGCNWSNMGNTGYYTNMGNTGYTGYGAYCCNGGNACNWSNATGT 250

Qy        353 AGCCAGGGCTTGCCGGGCCGCGA--TGGCCGCGACGGCCGCGACGGCGCGCCCGGGGCTC 410
Db        251 AYCCNGCNACNGCNGTNCCNCARATHAAYATHACNATHYTNAARGGNGARAARGGNGAYM 310

Qy        411 CGGGAGAGAAAGGCGAGGGCGGGAGGCCGGGACTGCCGGGACCTCGAGGGGACCCCGGGC 470
Db        311 GNGGNGAYMGNGGNYTNCARGGNAARTAYGGNAARACNGGNWSNGCNGGNGCNMGNGGNC 370

Qy        471 CGCGAGGAGAGGCGGGACCCGCGGGGCCCACCGGGCCTGCCGGGGAGTGCTCGGTGCCTC 530
Db        371 AYACNGGNCCNAARGGNCARAARGGNWSNATGGGNGCNCCNGGNGARMGNTGYAARWSNC 430

Qy        531 CGCGATCCGCCTTCAGCGCCAAGCGCTCCGAGAGCCGGGTGCCTCCGCCGTCTGACGCAC 590
Db        431 AYTAYGCNGCNTTYWSNGTNGGNMGNAARAARCCNATGCAYWSNAAYCAYTAYTAYCARA 490

Qy        591 CCTTGCCCTTCGACCGCGTGCTGGTGAACGAGCAGGGACATTACGACGCCGTCACCGGCA 650
Db        491 CNGTNATHTTYGAYACNGARTTYGTNAAYYTNTAYGAYCAYTTYAAYATGTTYACNGGNA 550

Qy        651 AGTTCACCTGCCAGGTGCCTGGGGTCTACTACTTCGCCGTCCATGC---CACCGTCTACC 707
Db        551 ARTTYTAYTGYTAYGTNCCNGGNYTNTAYTTYTTYWSNYTNAAYGTNCAYACNTGGAAYC 610

Qy        708 GGGCCAGCCTGCAGTTTGATCTGGTGAAGAATGGCGAATCCATTGCCTCTTTCTTCCAGT 767
Db        611 ARAARGARACNTAYYTNCAYATHATGAARAAYGARGARGARGTNGTNATHYTNTTYGCNC 670

Qy        768 TTTTCGGGGGGTGGCCCAAGCCAGCCTCGCTCTCGGGGGGGGCCATGGTGAGGCTGGAGC 827
Db        671 ARGTNGGNGAYMGNWSNATHATGCARWSNCARWSNYTNATGYTNGARYTNMGNGARCARG 730

Qy        828 CTGAGGACCAAGTGTGGGTGCAGGTGGGTGTGGGTGACTACATTGGCATCTATGCCAGCA 887
Db        731 AYCARGTNTGGGTNMGNYTNTAYAARGGNGARMGNGARAAYGCNATHTTYWSNGARGARY 790

Qy        888 TCAAGACAGACAGCACCTTCTCCGGATTTCTGGTGTA 924
Db        791 TNGAYACNTAYATHACNTTYWSNGGNTAYYTNGTNAA 827
```

```
RESULT 15
US-09-506-855-23
; Sequence 23, Application US/09506855
; Patent No. 6448221
; GENERAL INFORMATION:
;  APPLICANT: Sheppard, Paul O.
;  APPLICANT: Lasser, Gerald W.
;  APPLICANT: Bishop, Paul D.
;  TITLE OF INVENTION: INHIBITORS FOR USE IN HEMOSTASIS AND
;  TITLE OF INVENTION: IMMUNE FUNCTION
;  FILE REFERENCE: 99-12
;  CURRENT APPLICATION NUMBER: US/09/506,855
;  CURRENT FILING DATE: 2000-02-17
;  NUMBER OF SEQ ID NOS: 50
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 23
;   LENGTH: 843
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Degenerate nucleotide sequence encoding zsig37
;   OTHER INFORMATION: polypeptide
;   FEATURE:
;   NAME/KEY: variation
;   LOCATION: (1)...(843)
;   OTHER INFORMATION: Each N is independently any nucleotide.
US-09-506-855-23

  Query Match           5.6%;  Score 77.4;  DB 4;  Length 843;
  Best Local Similarity  30.0%;  Pred. No. 3e-07;
  Matches  209;  Conservative   99;  Mismatches  384;  Indels    5;  Gaps    2;

Qy       233 CCACTCCTCGTCCTGCTGCTCCTGGGCCTGGCGGCCGGCTCGCCCCCACTGGACGACAAC 292
Db       131 CNWSNCCNCCNGAYCAYGCNGARMGNGCNGARGARCARCAYGARAARTAYMGNCCNWSNC 190

Qy       293 AAGATCCCCAGCCTCTGCCCGGGGCACCCCGGCCTTCCAGGCACGCCGGGCCACCATGGC 352
Db       191 ARGAYCARGGNYTNCCNGCNWSNMGNTGYYTNMGNTGYTGYGAYCCNGGNACNWSNATGT 250

Qy       353 AGCCAGGGCTTGCCGGGCCGCGA--TGGCCGCGACGGCCGCGACGGCGCGCCCGGGGCTC 410
Db       251 AYCCNGCNACNGCNGTNCCNCARATHAAYATHACNATHYTNAARGGNGARAARGGNGAYM 310

Qy       411 CGGGAGAGAAAGGCGAGGGCGGGAGGCCGGGACTGCCGGGACCTCGAGGGGACCCCGGGC 470
Db       311 GNGGNGAYMGNGGNYTNCARGGNAARTAYGGNAARACNGGNWSNGCNGGNGCNMGNGGNC 370

Qy       471 CGCGAGGAGAGGCGGGACCCGCGGGGCCCACCGGGCCTGCCGGGGAGTGCTCGGTGCCTC 530
Db       371 AYACNGGNCCNAARGGNCARAARGGNWSNATGGGNGCNCCNGGNGARMGNTGYAARWSNC 430

Qy       531 CGCGATCCGCCTTCAGCGCCAAGCGCTCCGAGAGCCGGGTGCCTCCGCCGTCTGACGCAC 590
Db       431 AYTAYGCNGCNTTYWSNGTNGGNMGNAARAARCCNATGCAYWSNAAYCAYTAYTAYCARA 490

Qy       591 CCTTGCCCTTCGACCGCGTGCTGGTGAACGAGCAGGGACATTACGACGCCGTCACCGGCA 650
Db       491 CNGTNATHTTYGAYACNGARTTYGTNAAYYTNTAYGAYCAYTTYAAYATGTTYACNGGNA 550

Qy       651 AGTTCACCTGCCAGGTGCCTGGGGTCTACTACTTCGCCGTCCATGC---CACCGTCTACC 707
Db       551 ARTTYTAYTGYTAYGTNCCNGGNYTNTAYTTYTTYWSNYTNAAYGTNCAYACNTGGAAYC 610

Qy       708 GGGCCAGCCTGCAGTTTGATCTGGTGAAGAATGGCGAATCCATTGCCTCTTTCTTCCAGT 767
Db       611 ARAARGARACNTAYYTNCAYATHATGAARAAYGARGARGARGTNGTNATHYTNTTYGCNC 670

Qy       768 TTTTCGGGGGGTGGCCCAAGCCAGCCTCGCTCTCGGGGGGGGCCATGGTGAGGCTGGAGC 827
Db       671 ARGTNGGNGAYMGNWSNATHATGCARWSNCARWSNYTNATGYTNGARYTNMGNGARCARG 730

Qy       828 CTGAGGACCAAGTGTGGGTGCAGGTGGGTGTGGGTGACTACATTGGCATCTATGCCAGCA 887
Db       731 AYCARGTNTGGGTNMGNYTNTAYAARGGNGARMGNGARAAYGCNATHTTYWSNGARGARY 790

Qy       888 TCAAGACAGACAGCACCTTCTCCGGATTTCTGGTGTA 924
Db       791 TNGAYACNTAYATHACNTTYWSNGGNTAYYTNGTNAA 827

Search completed: June 22, 2003, 05:17:19
Job time : 106 secs
```

OM nucleic - nucleic search, using sw model

Run on: June 22, 2003, 04:43:36 ; Search time 225 Seconds
(without alignments)
8980.654 Million cell updates/sec

Title: US-09-943-851A-41
Perfect score: 1377
Sequence: 1 gactagttctcttggagtct.....aaaaaaaaaaaaaaaaaaa 1377

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1042519 seqs, 733713590 residues

Total number of hits satisfying chosen parameters: 2085038

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 1377 | 100.0 | 1377 | 9 | US-09-944-413-41 | Sequence 41, Appl |
| 2 | 1377 | 100.0 | 1377 | 9 | US-09-944-403-41 | Sequence 41, Appl |
| 3 | 1377 | 100.0 | 1377 | 9 | US-09-944-896-41 | Sequence 41, Appl |
| 4 | 1377 | 100.0 | 1377 | 9 | US-09-944-944-41 | Sequence 41, Appl |
| 5 | 1377 | 100.0 | 1377 | 9 | US-09-944-907-41 | Sequence 41, Appl |
| 6 | 1377 | 100.0 | 1377 | 9 | US-09-944-929-41 | Sequence 41, Appl |
| 7 | 1377 | 100.0 | 1377 | 9 | US-10-028-072-361 | Sequence 361, App |
| 8 | 1377 | 100.0 | 1377 | 9 | US-10-121-049-361 | Sequence 361, App |
| 9 | 1377 | 100.0 | 1377 | 9 | US-10-123-904-361 | Sequence 361, App |
| 10 | 1377 | 100.0 | 1377 | 9 | US-10-140-470-361 | Sequence 361, App |
| 11 | 1377 | 100.0 | 1377 | 9 | US-10-175-746-361 | Sequence 361, App |
| 12 | 1377 | 100.0 | 1377 | 9 | US-10-176-918-361 | Sequence 361, App |
| 13 | 1377 | 100.0 | 1377 | 9 | US-10-176-921-361 | Sequence 361, App |
| 14 | 1377 | 100.0 | 1377 | 9 | US-10-137-865-361 | Sequence 361, App |
| 15 | 1377 | 100.0 | 1377 | 9 | US-10-140-474-361 | Sequence 361, App |
| 16 | 1377 | 100.0 | 1377 | 9 | US-10-142-431-361 | Sequence 361, App |
| 17 | 1377 | 100.0 | 1377 | 9 | US-10-143-114-361 | Sequence 361, App |
| 18 | 1377 | 100.0 | 1377 | 9 | US-10-140-002-361 | Sequence 361, App |
| 19 | 1377 | 100.0 | 1377 | 9 | US-10-142-419-361 | Sequence 361, App |
| 20 | 1377 | 100.0 | 1377 | 9 | US-10-123-262-361 | Sequence 361, App |
| 21 | 1377 | 100.0 | 1377 | 9 | US-10-142-423-361 | Sequence 361, App |
| 22 | 1377 | 100.0 | 1377 | 9 | US-10-121-050-361 | Sequence 361, App |
| 23 | 1377 | 100.0 | 1377 | 9 | US-10-141-755-361 | Sequence 361, App |
| 24 | 1377 | 100.0 | 1377 | 9 | US-10-143-032-361 | Sequence 361, App |
| 25 | 1377 | 100.0 | 1377 | 9 | US-10-123-108-361 | Sequence 361, App |
| 26 | 1377 | 100.0 | 1377 | 9 | US-10-123-236-361 | Sequence 361, App |
| 27 | 1377 | 100.0 | 1377 | 9 | US-10-123-261-361 | Sequence 361, App |
| 28 | 1377 | 100.0 | 1377 | 9 | US-10-140-921-361 | Sequence 361, App |
| 29 | 1377 | 100.0 | 1377 | 9 | US-10-140-928-361 | Sequence 361, App |
| 30 | 1377 | 100.0 | 1377 | 9 | US-10-121-045-361 | Sequence 361, App |
| 31 | 1377 | 100.0 | 1377 | 9 | US-10-123-292-361 | Sequence 361, App |
| 32 | 1377 | 100.0 | 1377 | 9 | US-10-123-903-361 | Sequence 361, App |
| 33 | 1377 | 100.0 | 1377 | 9 | US-10-124-819-361 | Sequence 361, App |
| 34 | 1377 | 100.0 | 1377 | 9 | US-10-124-822-361 | Sequence 361, App |
| 35 | 1377 | 100.0 | 1377 | 9 | US-10-140-925-361 | Sequence 361, App |
| 36 | 1377 | 100.0 | 1377 | 9 | US-10-160-498-361 | Sequence 361, App |
| 37 | 1377 | 100.0 | 1377 | 9 | US-09-944-884-41 | Sequence 41, Appl |
| 38 | 1377 | 100.0 | 1377 | 9 | US-10-121-041-361 | Sequence 361, App |
| 39 | 1377 | 100.0 | 1377 | 9 | US-10-121-043-361 | Sequence 361, App |
| 40 | 1377 | 100.0 | 1377 | 9 | US-10-121-047-361 | Sequence 361, App |
| 41 | 1377 | 100.0 | 1377 | 9 | US-10-123-215-361 | Sequence 361, App |
| 42 | 1377 | 100.0 | 1377 | 9 | US-10-123-902-361 | Sequence 361, App |
| 43 | 1377 | 100.0 | 1377 | 9 | US-10-123-908-361 | Sequence 361, App |
| 44 | 1377 | 100.0 | 1377 | 9 | US-10-123-909-361 | Sequence 361, App |
| 45 | 1377 | 100.0 | 1377 | 9 | US-10-123-910-361 | Sequence 361, App |

ALIGNMENTS

RESULT 1
US-09-944-413-41
; Sequence 41, Application US/09944413
; Patent No. US20020156004A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin
; APPLICANT: Botstein,David
; APPLICANT: Eaton,Dan
; APPLICANT: Ferrara,Napoleone
; APPLICANT: Filvaroff,Ellen
; APPLICANT: Gerritsen,Mary
; APPLICANT: Goddard,Audrey
; APPLICANT: Godowski,Paul
; APPLICANT: Grimaldi,Christopher
; APPLICANT: Gurney,Austin
; APPLICANT: Hillan,Kenneth
; APPLICANT: Kljavin,Ivar
; APPLICANT: Napier,Mary
; APPLICANT: Roy,Margaret
; APPLICANT: Tumas,Daniel
; APPLICANT: Wood,William
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P2548P1C1
; CURRENT APPLICATION NUMBER: US/09/944,413
; CURRENT FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 09/866,028
; PRIOR FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: 60/067,411
; PRIOR FILING DATE: December 3, 1997
; PRIOR APPLICATION NUMBER: 60/069,334
; PRIOR FILING DATE: December 11, 1997
; PRIOR APPLICATION NUMBER: 60/069335
; PRIOR FILING DATE: December 11, 1997
; PRIOR APPLICATION NUMBER: 60/069,278
; PRIOR FILING DATE: December 11, 1997
; PRIOR APPLICATION NUMBER: 60/069,425
; PRIOR FILING DATE: December 12, 1997
; PRIOR APPLICATION NUMBER: 60/069,696
; PRIOR FILING DATE: December 16, 1997
; PRIOR APPLICATION NUMBER: 60/069,694
; PRIOR FILING DATE: December 16, 1997

; PRIOR APPLICATION NUMBER: 60/069,702
; PRIOR FILING DATE: December 16, 1997
; PRIOR APPLICATION NUMBER: 60/069,870
; PRIOR FILING DATE: December 17, 1997
; PRIOR APPLICATION NUMBER: 60/069,873
; PRIOR FILING DATE: December 17, 1997
; PRIOR APPLICATION NUMBER: 60/068,017
; PRIOR FILING DATE: December 18, 1997
; PRIOR APPLICATION NUMBER: 60/070,440
; PRIOR FILING DATE: January 5, 1998
; PRIOR APPLICATION NUMBER: 60/074,086
; PRIOR FILING DATE: February 9, 1998
; PRIOR APPLICATION NUMBER: 60/074,092
; PRIOR FILING DATE: February 9, 1998
; PRIOR APPLICATION NUMBER: 60/075,945
; PRIOR FILING DATE: February 25, 1998
; PRIOR APPLICATION NUMBER: 60/112,850
; PRIOR FILING DATE: December 16, 1998
; PRIOR APPLICATION NUMBER: 60/113,296
; PRIOR FILING DATE: December 22, 1998
; PRIOR APPLICATION NUMBER: 60/146,222
; PRIOR FILING DATE: July 28, 1999
; PRIOR APPLICATION NUMBER: PCT/US98/19330
; PRIOR FILING DATE: September 16, 1998
; PRIOR APPLICATION NUMBER: PCT/US98/25108
; PRIOR FILING DATE: December 1, 1998
; PRIOR APPLICATION NUMBER: 09/216,021
; PRIOR FILING DATE: December 16, 1998
; PRIOR APPLICATION NUMBER: 09/218,517
; PRIOR FILING DATE: December 22, 1998
; PRIOR APPLICATION NUMBER: 09/254,311
; PRIOR FILING DATE: March 3, 1999
; PRIOR APPLICATION NUMBER: PCT/US99/12252
; PRIOR FILING DATE: June 22, 1999
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: September 15, 1999
; PRIOR APPLICATION NUMBER: PCT/US99/28409
; PRIOR FILING DATE: No. US20020156004Alember 30, 1999
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: No. US20020156004Alember 30, 1999
; PRIOR APPLICATION NUMBER: PCT/US99/28301
; PRIOR FILING DATE: Decemberl, 1999
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: December 16, 1999
; PRIOR APPLICATION NUMBER: PCT/US00/03565
; PRIOR FILING DATE: February 11, 2000
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: February 22, 2000
; PRIOR APPLICATION NUMBER: PCT/US00/05841
; PRIOR FILING DATE: March 2, 2000
; PRIOR APPLICATION NUMBER: PCT/US00/08439
; PRIOR FILING DATE: March 30, 2000
; PRIOR APPLICATION NUMBER: PCT/US00/14042
; PRIOR FILING DATE: May 22, 2000
; PRIOR APPLICATION NUMBER: PCT/US00/20710
; PRIOR FILING DATE: July 28, 2000
; PRIOR APPLICATION NUMBER: PCT/US00/32678
; PRIOR FILING DATE: December 1, 2000
; PRIOR APPLICATION NUMBER: PCT/US01/06520
; PRIOR FILING DATE: February 28, 2001
; NUMBER OF SEQ ID NOS: 120
; SEQ ID NO 41
; LENGTH: 1377
; TYPE: DNA
; ORGANISM: Homo Sapien
US-09-944-413-41

Query Match 100.0%; Score 1377; DB 9; Length 1377;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1377; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 GACTAGTTCTCTTGGAGTCTGGGAGGAGGAAAGCGGAGCCGGCAGGGAGCGAACCAGGAC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GACTAGTTCTCTTGGAGTCTGGGAGGAGGAAAGCGGAGCCGGCAGGGAGCGAACCAGGAC 60

Qy         61 TGGGGTGACGGCAGGGCAGGGGGCGCCTGGCCGGGGAGAAGCGCGGGGGCTGGAGCACCA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TGGGGTGACGGCAGGGCAGGGGGCGCCTGGCCGGGGAGAAGCGCGGGGGCTGGAGCACCA 120

Qy        121 CCAACTGGAGGGTCCGGAGTAGCGAGCGCCCCGAAGGAGGCCATCGGGGAGCCGGGAGGG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CCAACTGGAGGGTCCGGAGTAGCGAGCGCCCCGAAGGAGGCCATCGGGGAGCCGGGAGGG 180

Qy        181 GGGACTGCGAGAGGACCCCGGCGTCCGGGCTCCCGGTGCCAGCGCTATGAGGCCACTCCT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGGACTGCGAGAGGACCCCGGCGTCCGGGCTCCCGGTGCCAGCGCTATGAGGCCACTCCT 240

Qy        241 CGTCCTGCTGCTCCTGGGCCTGGCGGCCGGCTCGCCCCCACTGGACGACAACAAGATCCC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CGTCCTGCTGCTCCTGGGCCTGGCGGCCGGCTCGCCCCCACTGGACGACAACAAGATCCC 300

Qy        301 CAGCCTCTGCCCGGGGCACCCCGGCCTTCCAGGCACGCCGGGCCACCATGGCAGCCAGGG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CAGCCTCTGCCCGGGGCACCCCGGCCTTCCAGGCACGCCGGGCCACCATGGCAGCCAGGG 360

Qy        361 CTTGCCGGGCCGCGATGGCCGCGACGGCCGCGACGGCGCGCCCGGGGCTCCGGGAGAGAA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CTTGCCGGGCCGCGATGGCCGCGACGGCCGCGACGGCGCGCCCGGGGCTCCGGGAGAGAA 420

Qy        421 AGGCGAGGGCGGGAGGCCGGGACTGCCGGGACCTCGAGGGGACCCCGGGCCGCGAGGAGA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AGGCGAGGGCGGGAGGCCGGGACTGCCGGGACCTCGAGGGGACCCCGGGCCGCGAGGAGA 480

Qy        481 GGCGGGACCCGCGGGGCCCACCGGGCCTGCCGGGGAGTGCTCGGTGCCTCCGCGATCCGC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GGCGGGACCCGCGGGGCCCACCGGGCCTGCCGGGGAGTGCTCGGTGCCTCCGCGATCCGC 540

Qy        541 CTTCAGCGCCAAGCGCTCCGAGAGCCGGGTGCCTCCGCCGTCTGACGCACCCTTGCCCTT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CTTCAGCGCCAAGCGCTCCGAGAGCCGGGTGCCTCCGCCGTCTGACGCACCCTTGCCCTT 600

Qy        601 CGACCGCGTGCTGGTGAACGAGCAGGGACATTACGACGCCGTCACCGGCAAGTTCACCTG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CGACCGCGTGCTGGTGAACGAGCAGGGACATTACGACGCCGTCACCGGCAAGTTCACCTG 660

Qy        661 CCAGGTGCCTGGGGTCTACTACTTCGCCGTCCATGCCACCGTCTACCGGGCCAGCCTGCA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CCAGGTGCCTGGGGTCTACTACTTCGCCGTCCATGCCACCGTCTACCGGGCCAGCCTGCA 720

Qy        721 GTTTGATCTGGTGAAGAATGGCGAATCCATTGCCTCTTTCTTCCAGTTTTTCGGGGGGTG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GTTTGATCTGGTGAAGAATGGCGAATCCATTGCCTCTTTCTTCCAGTTTTTCGGGGGGTG 780

Qy        781 GCCCAAGCCAGCCTCGCTCTCGGGGGGGGCCATGGTGAGGCTGGAGCCTGAGGACCAAGT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GCCCAAGCCAGCCTCGCTCTCGGGGGGGGCCATGGTGAGGCTGGAGCCTGAGGACCAAGT 840

Qy        841 GTGGGTGCAGGTGGGTGTGGGTGACTACATTGGCATCTATGCCAGCATCAAGACAGACAG 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GTGGGTGCAGGTGGGTGTGGGTGACTACATTGGCATCTATGCCAGCATCAAGACAGACAG 900

Qy        901 CACCTTCTCCGGATTTCTGGTGTACTCCGACTGGCACAGCTCCCCAGTCTTTGCTTAGTG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 CACCTTCTCCGGATTTCTGGTGTACTCCGACTGGCACAGCTCCCCAGTCTTTGCTTAGTG 960

Qy        961 CCCACTGCAAAGTGAGCTCATGCTCTCACTCCTAGAAGGAGGGTGTGAGGCTGACAACCA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 CCCACTGCAAAGTGAGCTCATGCTCTCACTCCTAGAAGGAGGGTGTGAGGCTGACAACCA 1020

Qy       1021 GGTCATCCAGGAGGGCTGGCCCCCCTGGAATATTGTGAATGACTAGGGAGGTGGGGTAGA 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GGTCATCCAGGAGGGCTGGCCCCCCTGGAATATTGTGAATGACTAGGGAGGTGGGGTAGA 1080

Qy       1081 GCACTCTCCGTCCTGCTGCTGGCAAGGAATGGGAACAGTGGCTGTCTGCGATCAGGTCTG 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 GCACTCTCCGTCCTGCTGCTGGCAAGGAATGGGAACAGTGGCTGTCTGCGATCAGGTCTG 1140
```

```
Qy       1141 GCAGCATGGGGCAGTGGCTGGATTTCTGCCCAAGACCAGAGGAGTGTGCTGTGCTGGCAA 1200
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 GCAGCATGGGGCAGTGGCTGGATTTCTGCCCAAGACCAGAGGAGTGTGCTGTGCTGGCAA 1200

Qy       1201 GTGTAAGTCCCCCAGTTGCTCTGGTCCAGGAGCCCACGGTGGGGTGCTCTCTTCCTGGTC 1260
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 GTGTAAGTCCCCCAGTTGCTCTGGTCCAGGAGCCCACGGTGGGGTGCTCTCTTCCTGGTC 1260

Qy       1261 CTCTGCTTCTCTGGATCCTCCCCACCCCCTCCTGCTCCTGGGGCCGGCCCTTTTCTCAGA 1320
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 CTCTGCTTCTCTGGATCCTCCCCACCCCCTCCTGCTCCTGGGGCCGGCCCTTTTCTCAGA 1320

Qy       1321 GATCACTCAATAAACCTAAGAACCCTCATAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1377
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 GATCACTCAATAAACCTAAGAACCCTCATAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1377
```

RESULT 2
US-09-944-403-41
; Sequence 41, Application US/09944403
; Patent No. US20020165143A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin
; APPLICANT: Botstein,David
; APPLICANT: Eaton,Dan
; APPLICANT: Ferrara,Napoleone
; APPLICANT: Filvaroff,Ellen
; APPLICANT: Gerritsen,Mary
; APPLICANT: Goddard,Audrey
; APPLICANT: Godowski,Paul
; APPLICANT: Grimaldi,Christopher
; APPLICANT: Gurney,Austin
; APPLICANT: Hillan,Kenneth
; APPLICANT: Kljavin,Ivar
; APPLICANT: Napier,Mary
; APPLICANT: Roy,Margaret
; APPLICANT: Tumas,Daniel
; APPLICANT: Wood,William
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P2548P1C1
; CURRENT APPLICATION NUMBER: US/09/944,403
; CURRENT FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 09/866,028
; PRIOR FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: 60/067,411
; PRIOR FILING DATE: December 3, 1997
; PRIOR APPLICATION NUMBER: 60/069,334
; PRIOR FILING DATE: December 11, 1997
; PRIOR APPLICATION NUMBER: 60/069335
; PRIOR FILING DATE: December 11, 1997
; PRIOR APPLICATION NUMBER: 60/069,278
; PRIOR FILING DATE: December 11, 1997
; PRIOR APPLICATION NUMBER: 60/069,425
; PRIOR FILING DATE: December 12, 1997
; PRIOR APPLICATION NUMBER: 60/069,696
; PRIOR FILING DATE: December 16, 1997
; PRIOR APPLICATION NUMBER: 60/069,694
; PRIOR FILING DATE: December 16, 1997
; PRIOR APPLICATION NUMBER: 60/069,702
; PRIOR FILING DATE: December 16, 1997
; PRIOR APPLICATION NUMBER: 60/069,870
; PRIOR FILING DATE: December 17, 1997
; PRIOR APPLICATION NUMBER: 60/069,873
; PRIOR FILING DATE: December 17, 1997
; PRIOR APPLICATION NUMBER: 60/068,017
; PRIOR FILING DATE: December 18, 1997
; PRIOR APPLICATION NUMBER: 60/070,440
; PRIOR FILING DATE: January 5, 1998
; PRIOR APPLICATION NUMBER: 60/074,086
; PRIOR FILING DATE: February 9, 1998
; PRIOR APPLICATION NUMBER: 60/074,092
; PRIOR FILING DATE: February 9, 1998
; PRIOR APPLICATION NUMBER: 60/075,945
; PRIOR FILING DATE: February 25, 1998
; PRIOR APPLICATION NUMBER: 60/112,850
; PRIOR FILING DATE: December 16, 1998
; PRIOR APPLICATION NUMBER: 60/113,296
; PRIOR FILING DATE: December 22, 1998
; PRIOR APPLICATION NUMBER: 60/146,222
; PRIOR FILING DATE: July 28, 1999
; PRIOR APPLICATION NUMBER: PCT/US98/19330
; PRIOR FILING DATE: September 16, 1998
; PRIOR APPLICATION NUMBER: PCT/US98/25108
; PRIOR FILING DATE: December 1, 1998
; PRIOR APPLICATION NUMBER: 09/216,021
; PRIOR FILING DATE: December 16, 1998
; PRIOR APPLICATION NUMBER: 09/218,517
; PRIOR FILING DATE: December 22, 1998
; PRIOR APPLICATION NUMBER: 09/254,311
; PRIOR FILING DATE: March 3, 1999
; PRIOR APPLICATION NUMBER: PCT/US99/12252
; PRIOR FILING DATE: June 22, 1999
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: September 15, 1999
; PRIOR APPLICATION NUMBER: PCT/US99/28409
; PRIOR FILING DATE: No. US20020165143Alember 30, 1999
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: No. US20020165143Alember 30, 1999
; PRIOR APPLICATION NUMBER: PCT/US99/28301
; PRIOR FILING DATE: Decemberl, 1999
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: December 16, 1999
; PRIOR APPLICATION NUMBER: PCT/US00/03565
; PRIOR FILING DATE: February 11, 2000
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: February 22, 2000
; PRIOR APPLICATION NUMBER: PCT/US00/05841
; PRIOR FILING DATE: March 2, 2000
; PRIOR APPLICATION NUMBER: PCT/US00/08439
; PRIOR FILING DATE: March 30, 2000
; PRIOR APPLICATION NUMBER: PCT/US00/14042
; PRIOR FILING DATE: May 22, 2000
; PRIOR APPLICATION NUMBER: PCT/US00/20710
; PRIOR FILING DATE: July 28, 2000
; PRIOR APPLICATION NUMBER: PCT/US00/32678
; PRIOR FILING DATE: December 1, 2000
; PRIOR APPLICATION NUMBER: PCT/US01/06520
; PRIOR FILING DATE: February 28, 2001
; NUMBER OF SEQ ID NOS: 120
; SEQ ID NO 41
; LENGTH: 1377
; TYPE: DNA
; ORGANISM: Homo Sapien
US-09-944-403-41

Query Match 100.0%; Score 1377; DB 9; Length 1377;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1377; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 GACTAGTTCTCTTGGAGTCTGGGAGGAGGAAAGCGGAGCCGGCAGGGAGCGAACCAGGAC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GACTAGTTCTCTTGGAGTCTGGGAGGAGGAAAGCGGAGCCGGCAGGGAGCGAACCAGGAC 60

Qy         61 TGGGGTGACGGCAGGGCAGGGGGCGCCTGGCCGGGGAGAAGCGCGGGGGCTGGAGCACCA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TGGGGTGACGGCAGGGCAGGGGGCGCCTGGCCGGGGAGAAGCGCGGGGGCTGGAGCACCA 120

Qy        121 CCAACTGGAGGGTCCGGAGTAGCGAGCGCCCCGAAGGAGGCCATCGGGGAGCCGGGAGGG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CCAACTGGAGGGTCCGGAGTAGCGAGCGCCCCGAAGGAGGCCATCGGGGAGCCGGGAGGG 180

Qy        181 GGGACTGCGAGAGGACCCCGGCGTCCGGGCTCCCGGTGCCAGCGCTATGAGGCCACTCCT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGGACTGCGAGAGGACCCCGGCGTCCGGGCTCCCGGTGCCAGCGCTATGAGGCCACTCCT 240
```

```
Qy       241 CGTCCTGCTGCTCCTGGGCCTGGCGGCCGGCTCGCCCCCACTGGACGACAACAAGATCCC 300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       241 CGTCCTGCTGCTCCTGGGCCTGGCGGCCGGCTCGCCCCCACTGGACGACAACAAGATCCC 300

Qy       301 CAGCCTCTGCCCGGGGCACCCCGGCCTTCCAGGCACGCCGGGCCACCATGGCAGCCAGGG 360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       301 CAGCCTCTGCCCGGGGCACCCCGGCCTTCCAGGCACGCCGGGCCACCATGGCAGCCAGGG 360

Qy       361 CTTGCCGGGCCGCGATGGCCGCGACGGCCGCGACGGCGCGCCCGGGGCTCCGGGAGAGAA 420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       361 CTTGCCGGGCCGCGATGGCCGCGACGGCCGCGACGGCGCGCCCGGGGCTCCGGGAGAGAA 420

Qy       421 AGGCGAGGGCGGGAGGCCGGGACTGCCGGGACCTCGAGGGGACCCCGGGCCGCGAGGAGA 480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       421 AGGCGAGGGCGGGAGGCCGGGACTGCCGGGACCTCGAGGGGACCCCGGGCCGCGAGGAGA 480

Qy       481 GGCGGGACCCGCGGGGCCCACCGGGCCTGCCGGGGAGTGCTCGGTGCCTCCGCGATCCGC 540
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       481 GGCGGGACCCGCGGGGCCCACCGGGCCTGCCGGGGAGTGCTCGGTGCCTCCGCGATCCGC 540

Qy       541 CTTCAGCGCCAAGCGCTCCGAGAGCCGGGTGCCTCCGCCGTCTGACGCACCCTTGCCCTT 600
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       541 CTTCAGCGCCAAGCGCTCCGAGAGCCGGGTGCCTCCGCCGTCTGACGCACCCTTGCCCTT 600

Qy       601 CGACCGCGTGCTGGTGAACGAGCAGGGACATTACGACGCCGTCACCGGCAAGTTCACCTG 660
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       601 CGACCGCGTGCTGGTGAACGAGCAGGGACATTACGACGCCGTCACCGGCAAGTTCACCTG 660

Qy       661 CCAGGTGCCTGGGGTCTACTACTTCGCCGTCCATGCCACCGTCTACCGGGCCAGCCTGCA 720
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       661 CCAGGTGCCTGGGGTCTACTACTTCGCCGTCCATGCCACCGTCTACCGGGCCAGCCTGCA 720

Qy       721 GTTTGATCTGGTGAAGAATGGCGAATCCATTGCCTCTTTCTTCCAGTTTTTCGGGGGGTG 780
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       721 GTTTGATCTGGTGAAGAATGGCGAATCCATTGCCTCTTTCTTCCAGTTTTTCGGGGGGTG 780

Qy       781 GCCCAAGCCAGCCTCGCTCTCGGGGGGGGCCATGGTGAGGCTGGAGCCTGAGGACCAAGT 840
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       781 GCCCAAGCCAGCCTCGCTCTCGGGGGGGGCCATGGTGAGGCTGGAGCCTGAGGACCAAGT 840

Qy       841 GTGGGTGCAGGTGGGTGTGGGTGACTACATTGGCATCTATGCCAGCATCAAGACAGACAG 900
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       841 GTGGGTGCAGGTGGGTGTGGGTGACTACATTGGCATCTATGCCAGCATCAAGACAGACAG 900

Qy       901 CACCTTCTCCGGATTTCTGGTGTACTCCGACTGGCACAGCTCCCCAGTCTTTGCTTAGTG 960
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       901 CACCTTCTCCGGATTTCTGGTGTACTCCGACTGGCACAGCTCCCCAGTCTTTGCTTAGTG 960

Qy       961 CCCACTGCAAAGTGAGCTCATGCTCTCACTCCTAGAAGGAGGGTGTGAGGCTGACAACCA 1020
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       961 CCCACTGCAAAGTGAGCTCATGCTCTCACTCCTAGAAGGAGGGTGTGAGGCTGACAACCA 1020

Qy      1021 GGTCATCCAGGAGGGCTGGCCCCCCCTGGAATATTGTGAATGACTAGGGAGGTGGGGTAGA 1080
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1021 GGTCATCCAGGAGGGCTGGCCCCCCCTGGAATATTGTGAATGACTAGGGAGGTGGGGTAGA 1080

Qy      1081 GCACTCTCCGTCCTGCTGCTGGCAAGGAATGGGAACAGTGGCTGTCTGCGATCAGGTCTG 1140
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1081 GCACTCTCCGTCCTGCTGCTGGCAAGGAATGGGAACAGTGGCTGTCTGCGATCAGGTCTG 1140

Qy      1141 GCAGCATGGGGCAGTGGCTGGATTTCTGCCCAAGACCAGAGGAGTGTGCTGTGCTGGCAA 1200
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1141 GCAGCATGGGGCAGTGGCTGGATTTCTGCCCAAGACCAGAGGAGTGTGCTGTGCTGGCAA 1200

Qy      1201 GTGTAAGTCCCCCAGTTGCTCTGGTCCAGGAGCCCACGGTGGGGTGCTCTCTTCCTGGTC 1260
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1201 GTGTAAGTCCCCCAGTTGCTCTGGTCCAGGAGCCCACGGTGGGGTGCTCTCTTCCTGGTC 1260

Qy      1261 CTCTGCTTCTCTGGATCCTCCCCACCCCCTCCTGCTCCTGGGGCCGGCCCTTTTCTCAGA 1320
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1261 CTCTGCTTCTCTGGATCCTCCCCACCCCCTCCTGCTCCTGGGGCCGGCCCTTTTCTCAGA 1320
```

```
Qy      1321 GATCACTCAATAAACCTAAGAACCCTCATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1377
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1321 GATCACTCAATAAACCTAAGAACCCTCATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1377
```

```
RESULT 3
US-09-944-896-41
; Sequence 41, Application US/09944896
; Patent No. US20020168715A1
; GENERAL INFORMATION:
;  APPLICANT: Baker, Kevin
;  APPLICANT:  Botstein,David
;  APPLICANT:  Eaton,Dan
;  APPLICANT:  Ferrara,Napoleone
;  APPLICANT:  Filvaroff,Ellen
;  APPLICANT:  Gerritsen,Mary
;  APPLICANT:  Goddard,Audrey
;  APPLICANT:  Godowski,Paul
;  APPLICANT:  Grimaldi,Christopher
;  APPLICANT:  Gurney,Austin
;  APPLICANT:  Hillan,Kenneth
;  APPLICANT:  Kljavin,Ivar
;  APPLICANT:  Napier,Mary
;  APPLICANT:  Roy,Margaret
;  APPLICANT:  Tumas,Daniel
;  APPLICANT:  Wood,William
;  TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;  TITLE OF INVENTION:  ACIDS ENCODING THE SAME
;  FILE REFERENCE: P2548P1C1
;  CURRENT APPLICATION NUMBER: US/09/944,896
;  CURRENT FILING DATE:  2001-08-31
;  PRIOR APPLICATION NUMBER: 09/866,028
;  PRIOR FILING DATE: 2001-05-25
;  PRIOR APPLICATION NUMBER: 60/069,334
;  PRIOR FILING DATE: December 11, 1997
;  PRIOR APPLICATION NUMBER: 60/069335
;  PRIOR FILING DATE: December 11, 1997
;  PRIOR APPLICATION NUMBER: 60/069,278
;  PRIOR FILING DATE: December 11, 1997
;  PRIOR APPLICATION NUMBER: 60/069,425
;  PRIOR FILING DATE: December 12, 1997
;  PRIOR APPLICATION NUMBER: 60/069,696
;  PRIOR FILING DATE: December 16, 1997
;  PRIOR APPLICATION NUMBER: 60/069,694
;  PRIOR FILING DATE: December 16, 1997
;  PRIOR APPLICATION NUMBER: 60/069,702
;  PRIOR FILING DATE: December 16, 1997
;  PRIOR APPLICATION NUMBER: 60/069,870
;  PRIOR FILING DATE: December 17, 1997
;  PRIOR APPLICATION NUMBER: 60/069,873
;  PRIOR FILING DATE: December 17, 1997
;  PRIOR APPLICATION NUMBER: 60/068,017
;  PRIOR FILING DATE: December 18, 1997
;  PRIOR APPLICATION NUMBER: 60/070,440
;  PRIOR FILING DATE: January 5, 1998
;  PRIOR APPLICATION NUMBER: 60/074,086
;  PRIOR FILING DATE: February 9, 1998
;  PRIOR APPLICATION NUMBER: 60/074,092
;  PRIOR FILING DATE: February 9, 1998
;  PRIOR APPLICATION NUMBER: 60/075,945
;  PRIOR FILING DATE: February 25, 1998
;  PRIOR APPLICATION NUMBER: 60/112,850
;  PRIOR FILING DATE: December 16, 1998
;  PRIOR APPLICATION NUMBER: 60/113,296
;  PRIOR FILING DATE: December 22, 1998
;  PRIOR APPLICATION NUMBER: 60/146,222
;  PRIOR FILING DATE: July 28, 1999
;  PRIOR APPLICATION NUMBER: PCT/US98/19330
;  PRIOR FILING DATE: September 16, 1998
;  PRIOR APPLICATION NUMBER: PCT/US98/25108
;  PRIOR FILING DATE: December 1, 1998
;  PRIOR APPLICATION NUMBER: 09/216,021
;  PRIOR FILING DATE: December 16, 1998
```

```
;  PRIOR APPLICATION NUMBER: 09/218,517
;  PRIOR FILING DATE: December 22, 1998
;  PRIOR APPLICATION NUMBER: 09/254,311
;  PRIOR FILING DATE: March 3, 1999
;  PRIOR APPLICATION NUMBER: PCT/US99/12252
;  PRIOR FILING DATE: June 22, 1999
;  PRIOR APPLICATION NUMBER: PCT/US99/21090
;  PRIOR FILING DATE: September 15, 1999
;  PRIOR APPLICATION NUMBER: PCT/US99/28409
;  PRIOR FILING DATE: No. US20020168715Alember 30, 1999
;  PRIOR APPLICATION NUMBER: PCT/US99/28313
;  PRIOR FILING DATE: No. US20020168715Alember 30, 1999
;  PRIOR APPLICATION NUMBER: PCT/US99/28301
;  PRIOR FILING DATE: Decemberl, 1999
;  PRIOR APPLICATION NUMBER: PCT/US99/30095
;  PRIOR FILING DATE: December 16, 1999
;  PRIOR APPLICATION NUMBER: PCT/US00/03565
;  PRIOR FILING DATE: February 11, 2000
;  PRIOR APPLICATION NUMBER: PCT/US00/04414
;  PRIOR FILING DATE: February 22, 2000
;  PRIOR APPLICATION NUMBER: PCT/US00/05841
;  PRIOR FILING DATE: March 2, 2000
;  PRIOR APPLICATION NUMBER: PCT/US00/08439
;  PRIOR FILING DATE: March 30, 2000
;  PRIOR APPLICATION NUMBER: PCT/US00/14042
;  PRIOR FILING DATE: May 22, 2000
;  PRIOR APPLICATION NUMBER: PCT/US00/20710
;  PRIOR FILING DATE: July 28, 2000
;  PRIOR APPLICATION NUMBER: PCT/US00/32678
;  PRIOR FILING DATE: December 1, 2000
;  PRIOR APPLICATION NUMBER: PCT/US01/06520
;  PRIOR FILING DATE: February 28, 2001
;  NUMBER OF SEQ ID NOS: 120
; SEQ ID NO 41
;   LENGTH: 1377
;   TYPE: DNA
;   ORGANISM: Homo Sapien
US-09-944-896-41

  Query Match             100.0%;  Score 1377;  DB 9;  Length 1377;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 1377;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GACTAGTTCTCTTGGAGTCTGGGAGGAGGAAAGCGGAGCCGGCAGGGAGCGAACCAGGAC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GACTAGTTCTCTTGGAGTCTGGGAGGAGGAAAGCGGAGCCGGCAGGGAGCGAACCAGGAC 60

Qy         61 TGGGGTGACGGCAGGGCAGGGGGCGCCTGGCCGGGGAGAAGCGCGGGGGCTGGAGCACCA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TGGGGTGACGGCAGGGCAGGGGGCGCCTGGCCGGGGAGAAGCGCGGGGGCTGGAGCACCA 120

Qy        121 CCAACTGGAGGGTCCGGAGTAGCGAGCGCCCCGAAGGAGGCCATCGGGGAGCCGGGAGGG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CCAACTGGAGGGTCCGGAGTAGCGAGCGCCCCGAAGGAGGCCATCGGGGAGCCGGGAGGG 180

Qy        181 GGGACTGCGAGAGGACCCCGGCGTCCGGGCTCCCGGTGCCAGCGCTATGAGGCCACTCCT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGGACTGCGAGAGGACCCCGGCGTCCGGGCTCCCGGTGCCAGCGCTATGAGGCCACTCCT 240

Qy        241 CGTCCTGCTGCTCCTGGGCCTGGCGGCCGGCTCGCCCCCACTGGACGACAACAAGATCCC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CGTCCTGCTGCTCCTGGGCCTGGCGGCCGGCTCGCCCCCACTGGACGACAACAAGATCCC 300

Qy        301 CAGCCTCTGCCCGGGGCACCCCGGCCTTCCAGGCACGCCGGGCCACCATGGCAGCCAGGG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CAGCCTCTGCCCGGGGCACCCCGGCCTTCCAGGCACGCCGGGCCACCATGGCAGCCAGGG 360

Qy        361 CTTGCCGGGCCGCGATGGCCGCGACGGCCGCGACGGCGCGCCCGGGGCTCCGGGAGAGAA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CTTGCCGGGCCGCGATGGCCGCGACGGCCGCGACGGCGCGCCCGGGGCTCCGGGAGAGAA 420

Qy        421 AGGCGAGGGCGGGAGGCCGGGACTGCCGGGACCTCGAGGGGACCCCGGGCCGCGAGGAGA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AGGCGAGGGCGGGAGGCCGGGACTGCCGGGACCTCGAGGGGACCCCGGGCCGCGAGGAGA 480

Qy        481 GGCGGGACCCGCGGGGCCCACCGGGCCTGCCGGGGAGTGCTCGGTGCCTCCGCGATCCGC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GGCGGGACCCGCGGGGCCCACCGGGCCTGCCGGGGAGTGCTCGGTGCCTCCGCGATCCGC 540

Qy        541 CTTCAGCGCCAAGCGCTCCGAGAGCCGGGTGCCTCCGCCGTCTGACGCACCCTTGCCCTT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CTTCAGCGCCAAGCGCTCCGAGAGCCGGGTGCCTCCGCCGTCTGACGCACCCTTGCCCTT 600

Qy        601 CGACCGCGTGCTGGTGAACGAGCAGGGACATTACGACGCCGTCACCGGCAAGTTCACCTG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CGACCGCGTGCTGGTGAACGAGCAGGGACATTACGACGCCGTCACCGGCAAGTTCACCTG 660

Qy        661 CCAGGTGCCTGGGGTCTACTACTTCGCCGTCCATGCCACCGTCTACCGGGCCAGCCTGCA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CCAGGTGCCTGGGGTCTACTACTTCGCCGTCCATGCCACCGTCTACCGGGCCAGCCTGCA 720

Qy        721 GTTTGATCTGGTGAAGAATGGCGAATCCATTGCCTCTTTCTTCCAGTTTTTCGGGGGGTG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GTTTGATCTGGTGAAGAATGGCGAATCCATTGCCTCTTTCTTCCAGTTTTTCGGGGGGTG 780

Qy        781 GCCCAAGCCAGCCTCGCTCTCGGGGGGGGCCATGGTGAGGCTGGAGCCTGAGGACCAAGT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GCCCAAGCCAGCCTCGCTCTCGGGGGGGGCCATGGTGAGGCTGGAGCCTGAGGACCAAGT 840

Qy        841 GTGGGTGCAGGTGGGTGTGGGTGACTACATTGGCATCTATGCCAGCATCAAGACAGACAG 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GTGGGTGCAGGTGGGTGTGGGTGACTACATTGGCATCTATGCCAGCATCAAGACAGACAG 900

Qy        901 CACCTTCTCCGGATTTCTGGTGTACTCCGACTGGCACAGCTCCCCAGTCTTTGCTTAGTG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 CACCTTCTCCGGATTTCTGGTGTACTCCGACTGGCACAGCTCCCCAGTCTTTGCTTAGTG 960

Qy        961 CCCACTGCAAAGTGAGCTCATGCTCTCACTCCTAGAAGGAGGGTGTGAGGCTGACAACCA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 CCCACTGCAAAGTGAGCTCATGCTCTCACTCCTAGAAGGAGGGTGTGAGGCTGACAACCA 1020

Qy       1021 GGTCATCCAGGAGGGCTGGCCCCCCTGGAATATTGTGAATGACTAGGGAGGTGGGGTAGA 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GGTCATCCAGGAGGGCTGGCCCCCCTGGAATATTGTGAATGACTAGGGAGGTGGGGTAGA 1080

Qy       1081 GCACTCTCCGTCCTGCTGCTGGCAAGGAATGGGAACAGTGGCTGTCTGCGATCAGGTCTG 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 GCACTCTCCGTCCTGCTGCTGGCAAGGAATGGGAACAGTGGCTGTCTGCGATCAGGTCTG 1140

Qy       1141 GCAGCATGGGGCAGTGGCTGGATTTCTGCCCAAGACCAGAGGAGTGTGCTGTGCTGGCAA 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 GCAGCATGGGGCAGTGGCTGGATTTCTGCCCAAGACCAGAGGAGTGTGCTGTGCTGGCAA 1200

Qy       1201 GTGTAAGTCCCCCAGTTGCTCTGGTCCAGGAGCCCACGGTGGGGTGCTCTCTTCCTGGTC 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 GTGTAAGTCCCCCAGTTGCTCTGGTCCAGGAGCCCACGGTGGGGTGCTCTCTTCCTGGTC 1260

Qy       1261 CTCTGCTTCTCTGGATCCTCCCCACCCCCTCCTGCTCCTGGGGCCGGCCCTTTTCTCAGA 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 CTCTGCTTCTCTGGATCCTCCCCACCCCCTCCTGCTCCTGGGGCCGGCCCTTTTCTCAGA 1320

Qy       1321 GATCACTCAATAAACCTAAGAACCCTCATAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1377
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 GATCACTCAATAAACCTAAGAACCCTCATAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1377

RESULT 4
US-09-944-944-41
; Sequence 41, Application US/09944944
; Patent No. US20020173463A1
; GENERAL INFORMATION:
;  APPLICANT: Baker, Kevin
;  APPLICANT:  Botstein,David
;  APPLICANT:  Eaton,Dan
;  APPLICANT:  Ferrara,Napoleone
;  APPLICANT:  Filvaroff,Ellen
```

; APPLICANT: Gerritsen,Mary
; APPLICANT: Goddard,Audrey
; APPLICANT: Godowski,Paul
; APPLICANT: Grimaldi,Christopher
; APPLICANT: Gurney,Austin
; APPLICANT: Hillan,Kenneth
; APPLICANT: Kljavin,Ivar
; APPLICANT: Napier,Mary
; APPLICANT: Roy,Margaret
; APPLICANT: Tumas,Daniel
; APPLICANT: Wood,William
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P2548P1C1
; CURRENT APPLICATION NUMBER: US/09/944,944
; CURRENT FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 09/866,028
; PRIOR FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: 60/067,411
; PRIOR FILING DATE: December 3, 1997
; PRIOR APPLICATION NUMBER: 60/069,334
; PRIOR FILING DATE: December 11, 1997
; PRIOR APPLICATION NUMBER: 60/069335
; PRIOR FILING DATE: December 11, 1997
; PRIOR APPLICATION NUMBER: 60/069,278
; PRIOR FILING DATE: December 11, 1997
; PRIOR APPLICATION NUMBER: 60/069,425
; PRIOR FILING DATE: December 12, 1997
; PRIOR APPLICATION NUMBER: 60/069,696
; PRIOR FILING DATE: December 16, 1997
; PRIOR APPLICATION NUMBER: 60/069,694
; PRIOR FILING DATE: December 16, 1997
; PRIOR APPLICATION NUMBER: 60/069,702
; PRIOR FILING DATE: December 16, 1997
; PRIOR APPLICATION NUMBER: 60/069,870
; PRIOR FILING DATE: December 17, 1997
; PRIOR APPLICATION NUMBER: 60/069,873
; PRIOR FILING DATE: December 17, 1997
; PRIOR APPLICATION NUMBER: 60/068,017
; PRIOR FILING DATE: December 18, 1997
; PRIOR APPLICATION NUMBER: 60/070,440
; PRIOR FILING DATE: January 5, 1998
; PRIOR APPLICATION NUMBER: 60/074,086
; PRIOR FILING DATE: February 9, 1998
; PRIOR APPLICATION NUMBER: 60/074,092
; PRIOR FILING DATE: February 9, 1998
; PRIOR APPLICATION NUMBER: 60/075,945
; PRIOR FILING DATE: February 25, 1998
; PRIOR APPLICATION NUMBER: 60/112,850
; PRIOR FILING DATE: December 16, 1998
; PRIOR APPLICATION NUMBER: 60/113,296
; PRIOR FILING DATE: December 22, 1998
; PRIOR APPLICATION NUMBER: 60/146,222
; PRIOR FILING DATE: July 28, 1999
; PRIOR APPLICATION NUMBER: PCT/US98/19330
; PRIOR FILING DATE: September 16, 1998
; PRIOR APPLICATION NUMBER: PCT/US98/25108
; PRIOR FILING DATE: December 1, 1998
; PRIOR APPLICATION NUMBER: 09/216,021
; PRIOR FILING DATE: December 16, 1998
; PRIOR APPLICATION NUMBER: 09/218,517
; PRIOR FILING DATE: December 22, 1998
; PRIOR APPLICATION NUMBER: 09/254,311
; PRIOR FILING DATE: March 3, 1999
; PRIOR APPLICATION NUMBER: PCT/US99/12252
; PRIOR FILING DATE: June 22, 1999
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: September 15, 1999
; PRIOR APPLICATION NUMBER: PCT/US99/28409
; PRIOR FILING DATE: No. US20020173463Alember 30, 1999
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: No. US20020173463Alember 30, 1999
; PRIOR APPLICATION NUMBER: PCT/US99/28301
; PRIOR FILING DATE: Decemberl, 1999
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: December 16, 1999
; PRIOR APPLICATION NUMBER: PCT/US00/03565
; PRIOR FILING DATE: February 11, 2000
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: February 22, 2000
; PRIOR APPLICATION NUMBER: PCT/US00/05841
; PRIOR FILING DATE: March 2, 2000
; PRIOR APPLICATION NUMBER: PCT/US00/08439
; PRIOR FILING DATE: March 30, 2000
; PRIOR APPLICATION NUMBER: PCT/US00/14042
; PRIOR FILING DATE: May 22, 2000
; PRIOR APPLICATION NUMBER: PCT/US00/20710
; PRIOR FILING DATE: July 28, 2000
; PRIOR APPLICATION NUMBER: PCT/US00/32678
; PRIOR FILING DATE: December 1, 2000
; PRIOR APPLICATION NUMBER: PCT/US01/06520
; PRIOR FILING DATE: February 28, 2001
; NUMBER OF SEQ ID NOS: 120
; SEQ ID NO 41
; LENGTH: 1377
; TYPE: DNA
; ORGANISM: Homo Sapien
US-09-944-944-41

Query Match 100.0%; Score 1377; DB 9; Length 1377;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1377; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 GACTAGTTCTCTTGGAGTCTGGGAGGAGGAAAGCGGAGCCGGCAGGGAGCGAACCAGGAC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GACTAGTTCTCTTGGAGTCTGGGAGGAGGAAAGCGGAGCCGGCAGGGAGCGAACCAGGAC 60

Qy         61 TGGGGTGACGGCAGGGCAGGGGGCGCCTGGCCGGGGAGAAGCGCGGGGGCTGGAGCACCA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TGGGGTGACGGCAGGGCAGGGGGCGCCTGGCCGGGGAGAAGCGCGGGGGCTGGAGCACCA 120

Qy        121 CCAACTGGAGGGTCCGGAGTAGCGAGCGCCCCGAAGGAGGCCATCGGGGAGCCGGGAGGG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CCAACTGGAGGGTCCGGAGTAGCGAGCGCCCCGAAGGAGGCCATCGGGGAGCCGGGAGGG 180

Qy        181 GGGACTGCGAGAGGACCCCGGCGTCCGGGCTCCCGGTGCCAGCGCTATGAGGCCACTCCT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGGACTGCGAGAGGACCCCGGCGTCCGGGCTCCCGGTGCCAGCGCTATGAGGCCACTCCT 240

Qy        241 CGTCCTGCTGCTCCTGGGCCTGGCGGCCGGCTCGCCCCCACTGGACGACAACAAGATCCC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CGTCCTGCTGCTCCTGGGCCTGGCGGCCGGCTCGCCCCCACTGGACGACAACAAGATCCC 300

Qy        301 CAGCCTCTGCCCGGGGCACCCCGGCCTTCCAGGCACGCCGGGCCACCATGGCAGCCAGGG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CAGCCTCTGCCCGGGGCACCCCGGCCTTCCAGGCACGCCGGGCCACCATGGCAGCCAGGG 360

Qy        361 CTTGCCGGGCCGCGATGGCCGCGACGGCCGCGACGGCGCGCCCGGGGCTCCGGGAGAGAA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CTTGCCGGGCCGCGATGGCCGCGACGGCCGCGACGGCGCGCCCGGGGCTCCGGGAGAGAA 420

Qy        421 AGGCGAGGGCGGGAGGCCGGGACTGCCGGGACCTCGAGGGGACCCCGGGCCGCGAGGAGA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AGGCGAGGGCGGGAGGCCGGGACTGCCGGGACCTCGAGGGGACCCCGGGCCGCGAGGAGA 480

Qy        481 GGCGGGACCCGCGGGGCCCACCGGGCCTGCCGGGGAGTGCTCGGTGCCTCCGCGATCCGC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GGCGGGACCCGCGGGGCCCACCGGGCCTGCCGGGGAGTGCTCGGTGCCTCCGCGATCCGC 540

Qy        541 CTTCAGCGCCAAGCGCTCCGAGAGCCGGGTGCCTCCGCCGTCTGACGCACCCTTGCCCTT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CTTCAGCGCCAAGCGCTCCGAGAGCCGGGTGCCTCCGCCGTCTGACGCACCCTTGCCCTT 600

Qy        601 CGACCGCGTGCTGGTGAACGAGCAGGGACATTACGACGCCGTCACCGGCAAGTTCACCTG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CGACCGCGTGCTGGTGAACGAGCAGGGACATTACGACGCCGTCACCGGCAAGTTCACCTG 660
```

```
Qy        661 CCAGGTGCCTGGGGTCTACTACTTCGCCGTCCATGCCACCGTCTACCGGGCCAGCCTGCA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CCAGGTGCCTGGGGTCTACTACTTCGCCGTCCATGCCACCGTCTACCGGGCCAGCCTGCA 720

Qy        721 GTTTGATCTGGTGAAGAATGGCGAATCCATTGCCTCTTTCTTCCAGTTTTTCGGGGGGGTG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GTTTGATCTGGTGAAGAATGGCGAATCCATTGCCTCTTTCTTCCAGTTTTTCGGGGGGGTG 780

Qy        781 GCCCAAGCCAGCCTCGCTCTCGGGGGGGGCCATGGTGAGGCTGGAGCCTGAGGACCAAGT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GCCCAAGCCAGCCTCGCTCTCGGGGGGGGCCATGGTGAGGCCTGGAGCCTGAGGACCAAGT 840

Qy        841 GTGGGTGCAGGTGGGTGTGGGTGACTACATTGGCATCTATGCCAGCATCAAGACAGACAG 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GTGGGTGCAGGTGGGTGTGGGTGACTACATTGGCATCTATGCCAGCATCAAGACAGACAG 900

Qy        901 CACCTTCTCCGGATTTCTGGTGTACTCCGACTGGCACAGCTCCCCAGTCTTTGCTTAGTG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 CACCTTCTCCGGATTTCTGGTGTACTCCGACTGGCACAGCTCCCCAGTCTTTGCTTAGTG 960

Qy        961 CCCACTGCAAAGTGAGCTCATGCTCTCACTCCTAGAAGGAGGGTGTGAGGCTGACAACCA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 CCCACTGCAAAGTGAGCTCATGCTCTCACTCCTAGAAGGAGGGTGTGAGGCTGACAACCA 1020

Qy       1021 GGTCATCCAGGAGGGCTGGCCCCCCTGGAATATTGTGAATGACTAGGGAGGTGGGGTAGA 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GGTCATCCAGGAGGGCTGGCCCCCCTGGAATATTGTGAATGACTAGGGAGGTGGGGTAGA 1080

Qy       1081 GCACTCTCCGTCCTGCTGCTGGCAAGGAATGGGAACAGTGGCTGTCTGCGATCAGGTCTG 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 GCACTCTCCGTCCTGCTGCTGGCAAGGAATGGGAACAGTGGCTGTCTGCGATCAGGTCTG 1140

Qy       1141 GCAGCATGGGGCAGTGGCTGGATTTCTGCCCAAGACCAGAGGAGTGTGCTGTGCTGGCAA 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 GCAGCATGGGGCAGTGGCTGGATTTCTGCCCAAGACCAGAGGAGTGTGCTGTGCTGGCAA 1200

Qy       1201 GTGTAAGTCCCCCAGTTGCTCTGGTCCAGGAGCCCACGGTGGGGTGCTCTCTTCCTGGTC 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 GTGTAAGTCCCCCAGTTGCTCTGGTCCAGGAGCCCACGGTGGGGTGCTCTCTTCCTGGTC 1260

Qy       1261 CTCTGCTTCTCTGGATCCTCCCCACCCCCTCCTGCTCCTGGGGCCGGCCCTTTTCTCAGA 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 CTCTGCTTCTCTGGATCCTCCCCACCCCCTCCTGCTCCTGGGGCCGGCCCTTTTCTCAGA 1320

Qy       1321 GATCACTCAATAAACCTAAGAACCCTCATAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1377
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 GATCACTCAATAAACCTAAGAACCCTCATAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1377

RESULT 5
US-09-944-907-41
; Sequence 41, Application US/09944907
; Publication No. US20020198147A1
; GENERAL INFORMATION:
;  APPLICANT: Baker, Kevin
;  APPLICANT:  Botstein,David
;  APPLICANT:  Eaton,Dan
;  APPLICANT:  Ferrara,Napoleone
;  APPLICANT:  Filvaroff,Ellen
;  APPLICANT:  Gerritsen,Mary
;  APPLICANT:  Goddard,Audrey
;  APPLICANT:  Godowski,Paul
;  APPLICANT:  Grimaldi,Christopher
;  APPLICANT:  Gurney,Austin
;  APPLICANT:  Hillan,Kenneth
;  APPLICANT:  Kljavin,Ivar
;  APPLICANT:  Napier,Mary
;  APPLICANT:  Roy,Margaret
;  APPLICANT:  Tumas,Daniel
;  APPLICANT:  Wood,William
;  TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;  TITLE OF INVENTION:  ACIDS ENCODING THE SAME
;  FILE REFERENCE: P2548P1C1
;  CURRENT APPLICATION NUMBER: US/09/944,907
;  CURRENT FILING DATE:  2001-08-31
;  PRIOR APPLICATION NUMBER: 09/866,028
;  PRIOR FILING DATE: 2001-05-25
;  NUMBER OF SEQ ID NOS: 120
; SEQ ID NO 41
;   LENGTH: 1377
;   TYPE: DNA
;   ORGANISM: Homo Sapien
US-09-944-907-41

  Query Match          100.0%;  Score 1377;  DB 9;  Length 1377;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches 1377;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GACTAGTTCTCTTGGAGTCTGGGAGGAGGAAAGCGGAGCCGGCAGGGAGCGAACCAGGAC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GACTAGTTCTCTTGGAGTCTGGGAGGAGGAAAGCGGAGCCGGCAGGGAGCGAACCAGGAC 60

Qy         61 TGGGGTGACGGCAGGGCAGGGGGCGCCTGGCCGGGGAGAAGCGCGGGGGCTGGAGCACCA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TGGGGTGACGGCAGGGCAGGGGGCGCCTGGCCGGGGAGAAGCGCGGGGGCTGGAGCACCA 120

Qy        121 CCAACTGGAGGGTCCGGAGTAGCGAGCGCCCCGAAGGAGGCCATCGGGGAGCCGGGAGGG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CCAACTGGAGGGTCCGGAGTAGCGAGCGCCCCGAAGGAGGCCATCGGGGAGCCGGGAGGG 180

Qy        181 GGGACTGCGAGAGGACCCCGGCGTCCGGGCTCCCGGTGCCAGCGCTATGAGGCCACTCCT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGGACTGCGAGAGGACCCCGGCGTCCGGGCTCCCGGTGCCAGCGCTATGAGGCCACTCCT 240

Qy        241 CGTCCTGCTGCTCCTGGGCCTGGCGGCCGGCTCGCCCCCACTGGACGACAACAAGATCCC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CGTCCTGCTGCTCCTGGGCCTGGCGGCCGGCTCGCCCCCACTGGACGACAACAAGATCCC 300

Qy        301 CAGCCTCTGCCCGGGGCACCCCGGCCTTCCAGGCACGCCGGGCCACCATGGCAGCCAGGG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CAGCCTCTGCCCGGGGCACCCCGGCCTTCCAGGCACGCCGGGCCACCATGGCAGCCAGGG 360

Qy        361 CTTGCCGGGCCGCGATGGCCGCGACGGCCGCGACGGCGCGCCCGGGGCTCCGGGAGAGAA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CTTGCCGGGCCGCGATGGCCGCGACGGCCGCGACGGCGCGCCCGGGGCTCCGGGAGAGAA 420

Qy        421 AGGCGAGGGCGGGAGGCCGGGACTGCCGGGACCTCGAGGGGACCCCGGGCCGCGAGGAGA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AGGCGAGGGCGGGAGGCCGGGACTGCCGGGACCTCGAGGGGACCCCGGGCCGCGAGGAGA 480

Qy        481 GGCGGGACCCGCGGGGCCCACCGGGCCTGCCGGGGAGTGCTCGGTGCCTCCGCGATCCGC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GGCGGGACCCGCGGGGCCCACCGGGCCTGCCGGGGAGTGCTCGGTGCCTCCGCGATCCGC 540

Qy        541 CTTCAGCGCCAAGCGCTCCGAGAGCCGGGTGCCTCCGCCGTCTGACGCACCCTTGCCCTT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CTTCAGCGCCAAGCGCTCCGAGAGCCGGGTGCCTCCGCCGTCTGACGCACCCTTGCCCTT 600

Qy        601 CGACCGCGTGCTGGTGAACGAGCAGGGACATTACGACGCCGTCACCGGCAAGTTCACCTG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CGACCGCGTGCTGGTGAACGAGCAGGGACATTACGACGCCGTCACCGGCAAGTTCACCTG 660

Qy        661 CCAGGTGCCTGGGGTCTACTACTTCGCCGTCCATGCCACCGTCTACCGGGCCAGCCTGCA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CCAGGTGCCTGGGGTCTACTACTTCGCCGTCCATGCCACCGTCTACCGGGCCAGCCTGCA 720

Qy        721 GTTTGATCTGGTGAAGAATGGCGAATCCATTGCCTCTTTCTTCCAGTTTTTCGGGGGGGTG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GTTTGATCTGGTGAAGAATGGCGAATCCATTGCCTCTTTCTTCCAGTTTTTCGGGGGGGTG 780

Qy        781 GCCCAAGCCAGCCTCGCTCTCGGGGGGGGCCATGGTGAGGCTGGAGCCTGAGGACCAAGT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GCCCAAGCCAGCCTCGCTCTCGGGGGGGGCCATGGTGAGGCTGGAGCCTGAGGACCAAGT 840

Qy        841 GTGGGTGCAGGTGGGTGTGGGTGACTACATTGGCATCTATGCCAGCATCAAGACAGACAG 900
```

```
Db        841 GTGGGTGCAGGTGGGTGTGGGTGACTACATTGGCATCTATGCCAGCATCAAGACAGACAG 900
Qy        901 CACCTTCTCCGGATTTCTGGTGTACTCCGACTGGCACAGCTCCCCAGTCTTTGCTTAGTG 960
Db        901 CACCTTCTCCGGATTTCTGGTGTACTCCGACTGGCACAGCTCCCCAGTCTTTGCTTAGTG 960
Qy        961 CCCACTGCAAAGTGAGCTCATGCTCTCACTCCTAGAAGGAGGGTGTGAGGCTGACAACCA 1020
Db        961 CCCACTGCAAAGTGAGCTCATGCTCTCACTCCTAGAAGGAGGGTGTGAGGCTGACAACCA 1020
Qy       1021 GGTCATCCAGGAGGGCTGGCCCCCCCTGGAATATTGTGAATGACTAGGGAGGTGGGGTAGA 1080
Db       1021 GGTCATCCAGGAGGGCTGGCCCCCCCTGGAATATTGTGAATGACTAGGGAGGTGGGGTAGA 1080
Qy       1081 GCACTCTCCGTCCTGCTGCTGGCAAGGAATGGGAACAGTGGCTGTCTGCGATCAGGTCTG 1140
Db       1081 GCACTCTCCGTCCTGCTGCTGGCAAGGAATGGGAACAGTGGCTGTCTGCGATCAGGTCTG 1140
Qy       1141 GCAGCATGGGGCAGTGGCTGGATTTCTGCCCAAGACCAGAGGAGTGTGCTGTGCTGGCAA 1200
Db       1141 GCAGCATGGGGCAGTGGCTGGATTTCTGCCCAAGACCAGAGGAGTGTGCTGTGCTGGCAA 1200
Qy       1201 GTGTAAGTCCCCCAGTTGCTCTGGTCCAGGAGCCCACGGTGGGGTGCTCTCTTCCTGGTC 1260
Db       1201 GTGTAAGTCCCCCAGTTGCTCTGGTCCAGGAGCCCACGGTGGGGTGCTCTCTTCCTGGTC 1260
Qy       1261 CTCTGCTTCTCTGGATCCTCCCCACCCCCTCCTGCTCCTGGGGCCGGCCCTTTTCTCAGA 1320
Db       1261 CTCTGCTTCTCTGGATCCTCCCCACCCCCTCCTGCTCCTGGGGCCGGCCCTTTTCTCAGA 1320
Qy       1321 GATCACTCAATAAACCTAAGAACCCTCATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1377
Db       1321 GATCACTCAATAAACCTAAGAACCCTCATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1377
```

RESULT 6
US-09-944-929-41
; Sequence 41, Application US/09944929
; Publication No. US20020197612A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin
; APPLICANT: Botstein,David
; APPLICANT: Eaton,Dan
; APPLICANT: Ferrara,Napoleone
; APPLICANT: Filvaroff,Ellen
; APPLICANT: Gerritsen,Mary
; APPLICANT: Goddard,Audrey
; APPLICANT: Godowski,Paul
; APPLICANT: Grimaldi,Christopher
; APPLICANT: Gurney,Austin
; APPLICANT: Hillan,Kenneth
; APPLICANT: Kljavin,Ivar
; APPLICANT: Napier,Mary
; APPLICANT: Roy,Margaret
; APPLICANT: Tumas,Daniel
; APPLICANT: Wood,William
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P2548P1C1
; CURRENT APPLICATION NUMBER: US/09/944,929
; CURRENT FILING DATE: 2001-08-31
; PRIOR APPLICATION NUMBER: 09/866,028
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 120
; SEQ ID NO 41
; LENGTH: 1377
; TYPE: DNA
; ORGANISM: Homo Sapien
US-09-944-929-41

Query Match 100.0%; Score 1377; DB 9; Length 1377;
Best Local Similarity 100.0%; Pred. No. 0;

Matches 1377; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 GACTAGTTCTCTTGGAGTCTGGGAGGAGGAAAGCGGAGCCGGCAGGGAGCGAACCAGGAC 60
Db          1 GACTAGTTCTCTTGGAGTCTGGGAGGAGGAAAGCGGAGCCGGCAGGGAGCGAACCAGGAC 60
Qy         61 TGGGGTGACGGCAGGGCAGGGGGCGCCTGGCCGGGGAGAAGCGCGGGGGCTGGAGCACCA 120
Db         61 TGGGGTGACGGCAGGGCAGGGGGCGCCTGGCCGGGGAGAAGCGCGGGGGCTGGAGCACCA 120
Qy        121 CCAACTGGAGGGTCCGGAGTAGCGAGCGCCCCGAAGGAGGCCATCGGGGAGCCGGGAGGG 180
Db        121 CCAACTGGAGGGTCCGGAGTAGCGAGCGCCCCGAAGGAGGCCATCGGGGAGCCGGGAGGG 180
Qy        181 GGGACTGCGAGAGGACCCCGGCGTCCGGGCTCCCGGTGCCAGCGCTATGAGGCCACTCCT 240
Db        181 GGGACTGCGAGAGGACCCCGGCGTCCGGGCTCCCGGTGCCAGCGCTATGAGGCCACTCCT 240
Qy        241 CGTCCTGCTGCTCCTGGGCCTGGCGGCCGGCTCGCCCCCACTGGACGACAACAAGATCCC 300
Db        241 CGTCCTGCTGCTCCTGGGCCTGGCGGCCGGCTCGCCCCCACTGGACGACAACAAGATCCC 300
Qy        301 CAGCCTCTGCCCGGGGCACCCCGGCCTTCCAGGCACGCCGGGCCACCATGGCAGCCAGGG 360
Db        301 CAGCCTCTGCCCGGGGCACCCCGGCCTTCCAGGCACGCCGGGCCACCATGGCAGCCAGGG 360
Qy        361 CTTGCCGGGCCGCGATGGCCGCGACGGCCGCGACGGCGCGCCCGGGGCTCCGGGAGAGAA 420
Db        361 CTTGCCGGGCCGCGATGGCCGCGACGGCCGCGACGGCGCGCCCGGGGCTCCGGGAGAGAA 420
Qy        421 AGGCGAGGGCGGGAGGCCGGGACTGCCGGGACCTCGAGGGGACCCCGGGCCGCGAGGAGA 480
Db        421 AGGCGAGGGCGGGAGGCCGGGACTGCCGGGACCTCGAGGGGACCCCGGGCCGCGAGGAGA 480
Qy        481 GGCGGGACCCGCGGGGCCCACCGGGCCTGCCGGGGAGTGCTCGGTGCCTCCGCGATCCGC 540
Db        481 GGCGGGACCCGCGGGGCCCACCGGGCCTGCCGGGGAGTGCTCGGTGCCTCCGCGATCCGC 540
Qy        541 CTTCAGCGCCAAGCGCTCCGAGAGCCGGGTGCCTCCGCCGTCTGACGCACCCTTGCCCTT 600
Db        541 CTTCAGCGCCAAGCGCTCCGAGAGCCGGGTGCCTCCGCCGTCTGACGCACCCTTGCCCTT 600
Qy        601 CGACCGCGTGCTGGTGAACGAGCAGGGACATTACGACGCCGTCACCGGCAAGTTCACCTG 660
Db        601 CGACCGCGTGCTGGTGAACGAGCAGGGACATTACGACGCCGTCACCGGCAAGTTCACCTG 660
Qy        661 CCAGGTGCCTGGGGTCTACTACTTCGCCGTCCATGCCACCGTCTACCGGGCCAGCCTGCA 720
Db        661 CCAGGTGCCTGGGGTCTACTACTTCGCCGTCCATGCCACCGTCTACCGGGCCAGCCTGCA 720
Qy        721 GTTTGATCTGGTGAAGAATGGCGAATCCATTGCCTCTTTCTTCCAGTTTTTCGGGGGGTG 780
Db        721 GTTTGATCTGGTGAAGAATGGCGAATCCATTGCCTCTTTCTTCCAGTTTTTCGGGGGGTG 780
Qy        781 GCCCAAGCCAGCCTCGCTCTCGGGGGGGGCCATGGTGAGGCTGGAGCCTGAGGACCAAGT 840
Db        781 GCCCAAGCCAGCCTCGCTCTCGGGGGGGGCCATGGTGAGGCTGGAGCCTGAGGACCAAGT 840
Qy        841 GTGGGTGCAGGTGGGTGTGGGTGACTACATTGGCATCTATGCCAGCATCAAGACAGACAG 900
Db        841 GTGGGTGCAGGTGGGTGTGGGTGACTACATTGGCATCTATGCCAGCATCAAGACAGACAG 900
Qy        901 CACCTTCTCCGGATTTCTGGTGTACTCCGACTGGCACAGCTCCCCAGTCTTTGCTTAGTG 960
Db        901 CACCTTCTCCGGATTTCTGGTGTACTCCGACTGGCACAGCTCCCCAGTCTTTGCTTAGTG 960
Qy        961 CCCACTGCAAAGTGAGCTCATGCTCTCACTCCTAGAAGGAGGGTGTGAGGCTGACAACCA 1020
Db        961 CCCACTGCAAAGTGAGCTCATGCTCTCACTCCTAGAAGGAGGGTGTGAGGCTGACAACCA 1020
Qy       1021 GGTCATCCAGGAGGGCTGGCCCCCCCTGGAATATTGTGAATGACTAGGGAGGTGGGGTAGA 1080
Db       1021 GGTCATCCAGGAGGGCTGGCCCCCCCTGGAATATTGTGAATGACTAGGGAGGTGGGGTAGA 1080
```

```
Qy       1081 GCACTCTCCGTCCTGCTGCTGGCAAGGAATGGGAACAGTGGCTGTCTGCGATCAGGTCTG 1140
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 GCACTCTCCGTCCTGCTGCTGGCAAGGAATGGGAACAGTGGCTGTCTGCGATCAGGTCTG 1140

Qy       1141 GCAGCATGGGGCAGTGGCTGGATTTCTGCCCAAGACCAGAGGAGTGTGCTGTGCTGGCAA 1200
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 GCAGCATGGGGCAGTGGCTGGATTTCTGCCCAAGAGCAGAGGAGTGTGCTGTGCTGGCAA 1200

Qy       1201 GTGTAAGTCCCCCAGTTGCTCTGGTCCAGGAGCCCACGGTGGGGTGCTCTCTTCCTGGTC 1260
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 GTGTAAGTCCCCCAGTTGCTCTGGTCCAGGAGCCCACGGTGGGGTGCTCTCTTCCTGGTC 1260

Qy       1261 CTCTGCTTCTCTGGATCCTCCCCACCCCCTCCTGCTCCTGGGGCCGGCCCTTTTCTCAGA 1320
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 CTCTGCTTCTCTGGATCCTCCCCACCCCCTCCTGCTCCTGGGGCCGGCCCTTTTCTCAGA 1320

Qy       1321 GATCACTCAATAAACCTAAGAACCCTCATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1377
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 GATCACTCAATAAACCTAAGAACCCTCATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1377
```

RESULT 7
US-10-028-072-361
; Sequence 361, Application US/10028072
; Publication No. US20030004311A1
; GENERAL INFORMATION:
; APPLICANT: Baker,Kevin P.
; APPLICANT: Beresini,Maureen
; APPLICANT: DeForge,Laura
; APPLICANT: Desnoyers,Luc
; APPLICANT: Filvaroff,Ellen
; APPLICANT: Gao,Wei-Qiang
; APPLICANT: Gerritsen,Mary E.
; APPLICANT: Goddard,Audrey
; APPLICANT: Godowski,Paul J.
; APPLICANT: Gurney,Austin L.
; APPLICANT: Sherwood,Steven
; APPLICANT: Smith,Victoria
; APPLICANT: Stewart,Timothy A.
; APPLICANT: Tumas,Daniel
; APPLICANT: Watanabe,Colin K
; APPLICANT: Wood,William
; APPLICANT: Zhang
; TITLE OF INVENTION:
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/028,072
; CURRENT FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/049911
; PRIOR FILING DATE: 1997-06-18
; PRIOR APPLICATION NUMBER: 60/056974
; PRIOR FILING DATE: 1997-08-26
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059115
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059117
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059122
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059184
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059352
; PRIOR FILING DATE: 1997-09-19
; PRIOR APPLICATION NUMBER: 60/059588
; PRIOR FILING DATE: 1997-09-19
; PRIOR APPLICATION NUMBER: 60/059836
; PRIOR FILING DATE: 1997-09-24
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/062285
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/062287
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/062814
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/062816
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063045
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063082
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: 60/063127
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063327
; PRIOR FILING DATE: 1997-10-27
; PRIOR APPLICATION NUMBER: 60/063329
; PRIOR FILING DATE: 1997-10-27
; PRIOR APPLICATION NUMBER: 60/063550
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063561
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063704
; PRIOR FILING DATE: 1997-10-29
; PRIOR APPLICATION NUMBER: 60/063733
; PRIOR FILING DATE: 1997-10-29
; PRIOR APPLICATION NUMBER: 60/063735
; PRIOR FILING DATE: 1997-10-29
; PRIOR APPLICATION NUMBER: 60/063738
; PRIOR FILING DATE: 1997-10-29
; PRIOR APPLICATION NUMBER: 60/063755
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/064248
; PRIOR FILING DATE: 1997-11-03
; PRIOR APPLICATION NUMBER: 60/064809
; PRIOR FILING DATE: 1997-11-07
; PRIOR APPLICATION NUMBER: 60/065186
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/065846
; PRIOR FILING DATE: 1997-11-17
; PRIOR APPLICATION NUMBER: 60/066364
; PRIOR FILING DATE: 1997-11-21
; PRIOR APPLICATION NUMBER: 60/066453
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/066511
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/066770
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/069212
; PRIOR FILING DATE: 1997-12-11
; PRIOR APPLICATION NUMBER: 60/069278
; PRIOR FILING DATE: 1997-12-11
; PRIOR APPLICATION NUMBER: 60/069334
; PRIOR FILING DATE: 1997-12-11
; PRIOR APPLICATION NUMBER: 60/069694
; PRIOR FILING DATE: 1997-12-16
; PRIOR APPLICATION NUMBER: 60/072320
; PRIOR FILING DATE: 1998-01-23
; PRIOR APPLICATION NUMBER: 60/073612
; PRIOR FILING DATE: 1998-02-04
; PRIOR APPLICATION NUMBER: 60/074086
; PRIOR FILING DATE: 1998-02-09
; PRIOR APPLICATION NUMBER: 60/074092
; PRIOR FILING DATE: 1998-02-09
; PRIOR APPLICATION NUMBER: 60/077791
; PRIOR FILING DATE: 1998-03-12
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/079294
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: 60/079663
; PRIOR FILING DATE: 1998-02-27
; PRIOR APPLICATION NUMBER: 60/079728
; PRIOR FILING DATE: 1998-03-27

; PRIOR APPLICATION NUMBER: 60/080165
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/081203
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: 60/081229
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: 60/081695
; PRIOR FILING DATE: 1998-04-14
; PRIOR APPLICATION NUMBER: 60/081817
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081818
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/082999
; PRIOR FILING DATE: 1998-04-24
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/083545
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084627
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084637
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/085149
; PRIOR FILING DATE: 1998-05-12
; PRIOR APPLICATION NUMBER: 60/085323
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085338
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085339
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085579
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085697
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085704
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/086414
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/086430
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/087106
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: 60/088026
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088730
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088741
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088810
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088858
; PRIOR FILING DATE: 19/98-06-11
; PRIOR APPLICATION NUMBER: 60/089532
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089599
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089907
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089947
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/090349
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090429
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090445
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090538
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090863
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/091360
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091519
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091982
; PRIOR FILING DATE: 1998-07-07

Query Match 100.0%; Score 1377; DB 9; Length 1377;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1377; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 GACTAGTTCTCTTGGAGTCTGGGAGGAGGAAAGCGGAGCCGGCAGGGAGCGAACCAGGAC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GACTAGTTCTCTTGGAGTCTGGGAGGAGGAAAGCGGAGCCGGCAGGGAGCGAACCAGGAC 60

Qy         61 TGGGGTGACGGCAGGGCAGGGGGCGCCTGGCCGGGGAGAAGCGCGGGGGCTGGAGCACCA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TGGGGTGACGGCAGGGCAGGGGGCGCCTGGCCGGGGAGAAGCGCGGGGGCTGGAGCACCA 120

Qy        121 CCAACTGGAGGGTCCGGAGTAGCGAGCGCCCCGAAGGAGGCCATCGGGGAGCCGGGAGGG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CCAACTGGAGGGTCCGGAGTAGCGAGCGCCCCGAAGGAGGCCATCGGGGAGCCGGGAGGG 180

Qy        181 GGGACTGCGAGAGGACCCCGGCGTCCGGGCTCCCGGTGCCAGCGCTATGAGGCCACTCCT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGGACTGCGAGAGGACCCCGGCGTCCGGGCTCCCGGTGCCAGCGCTATGAGGCCACTCCT 240

Qy        241 CGTCCTGCTGCTCCTGGGCCTGGCGGCCGGCTCGCCCCCACTGGACGACAACAAGATCCC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CGTCCTGCTGCTCCTGGGCCTGGCGGCCGGCTCGCCCCCACTGGACGACAACAAGATCCC 300

Qy        301 CAGCCTCTGCCCGGGGCACCCCGGCCTTCCAGGCACGCCGGGCCACCATGGCAGCCAGGG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CAGCCTCTGCCCGGGGCACCCCGGCCTTCCAGGCACGCCGGGCCACCATGGCAGCCAGGG 360

Qy        361 CTTGCCGGGCCGCGATGGCCGCGACGGCCGCGACGGCGCGCCCGGGGCTCCGGGAGAGAA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CTTGCCGGGCCGCGATGGCCGCGACGGCCGCGACGGCGCGCCCGGGGCTCCGGGAGAGAA 420

Qy        421 AGGCGAGGGCGGGAGGCCGGGACTGCCGGGACCTCGAGGGGACCCCGGGCCGCGAGGAGA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AGGCGAGGGCGGGAGGCCGGGACTGCCGGGACCTCGAGGGGACCCCGGGCCGCGAGGAGA 480

Qy        481 GGCGGGACCCGCGGGGCCCACCGGGCCTGCCGGGGAGTGCTCGGTGCCTCCGCGATCCGC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GGCGGGACCCGCGGGGCCCACCGGGCCTGCCGGGGAGTGCTCGGTGCCTCCGCGATCCGC 540

Qy        541 CTTCAGCGCCAAGCGCTCCGAGAGCCGGGTGCCTCCGCCGTCTGACGCACCCTTGCCCTT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CTTCAGCGCCAAGCGCTCCGAGAGCCGGGTGCCTCCGCCGTCTGACGCACCCTTGCCCTT 600

Qy        601 CGACCGCGTGCTGGTGAACGAGCAGGGACATTACGACGCCGTCACCGGCAAGTTCACCTG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CGACCGCGTGCTGGTGAACGAGCAGGGACATTACGACGCCGTCACCGGCAAGTTCACCTG 660

Qy        661 CCAGGTGCCTGGGGTCTACTACTTCGCCGTCCATGCCACCGTCTACCGGGCCAGCCTGCA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CCAGGTGCCTGGGGTCTACTACTTCGCCGTCCATGCCACCGTCTACCGGGCCAGCCTGCA 720

Qy        721 GTTTGATCTGGTGAAGAATGGCGAATCCATTGCCTCTTTCTTCCAGTTTTTCGGGGGGTG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GTTTGATCTGGTGAAGAATGGCGAATCCATTGCCTCTTTCTTCCAGTTTTTCGGGGGGTG 780

Qy        781 GCCCAAGCCAGCCTCGCTCTCGGGGGGGGCCATGGTGAGGCTGGAGCCTGAGGACCAAGT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GCCCAAGCCAGCCTCGCTCTCGGGGGGGGCCATGGTGAGGCTGGAGCCTGAGGACCAAGT 840

Qy        841 GTGGGTGCAGGTGGGTGTGGGTGACTACATTGGCATCTATGCCAGCATCAAGACAGACAG 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GTGGGTGCAGGTGGGTGTGGGTGACTACATTGGCATCTATGCCAGCATCAAGACAGACAG 900

Qy        901 CACCTTCTCCGGATTTCTGGTGTACTCCGACTGGCACAGCTCCCCAGTCTTTGCTTAGTG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 CACCTTCTCCGGATTTCTGGTGTACTCCGACTGGCACAGCTCCCCAGTCTTTGCTTAGTG 960
```

```
Qy        961 CCCACTGCAAAGTGAGCTCATGCTCTCACTCCTAGAAGGAGGGTGTGAGGCTGACAACCA 1020
Db        961 CCCACTGCAAAGTGAGCTCATGCTCTCACTCCTAGAAGGAGGGTGTGAGGCTGACAACCA 1020

Qy       1021 GGTCATCCAGGAGGGCTGGCCCCCCCTGGAATATTGTGAATGACTAGGGAGGTGGGGTAGA 1080
Db       1021 GGTCATCCAGGAGGGCTGGCCCCCCCTGGAATATTGTGAATGACTAGGGAGGTGGGGTAGA 1080

Qy       1081 GCACTCTCCGTCCTGCTGCTGGCAAGGAATGGGAACAGTGGCTGTCTGCGATCAGGTCTG 1140
Db       1081 GCACTCTCCGTCCTGCTGCTGGCAAGGAATGGGAACAGTGGCTGTCTGCGATCAGGTCTG 1140

Qy       1141 GCAGCATGGGGCAGTGGCTGGATTTCTGCCCAAGACCAGAGGAGTGTGCTGTGCTGGCAA 1200
Db       1141 GCAGCATGGGGCAGTGGCTGGATTTCTGCCCAAGACCAGAGGAGTGTGCTGTGCTGGCAA 1200

Qy       1201 GTGTAAGTCCCCCAGTTGCTCTGGTCCAGGAGCCCACGGTGGGGTGCTCTCTTCCTGGTC 1260
Db       1201 GTGTAAGTCCCCCAGTTGCTCTGGTCCAGGAGCCCACGGTGGGGTGCTCTCTTCCTGGTC 1260

Qy       1261 CTCTGCTTCTCTGGATCCTCCCCACCCCCTCCTGCTCCTGGGGCCGGCCCTTTTCTCAGA 1320
Db       1261 CTCTGCTTCTCTGGATCCTCCCCACCCCCTCCTGCTCCTGGGGCCGGCCCTTTTCTCAGA 1320

Qy       1321 GATCACTCAATAAACCTAAGAACCCTCATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1377
Db       1321 GATCACTCAATAAACCTAAGAACCCTCATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1377
```

RESULT 8
US-10-121-049-361
; Sequence 361, Application US/10121049
; Publication No. US20030022239A1
; GENERAL INFORMATION:
; APPLICANT: Baker,Kevin P.
; APPLICANT: Beresini,Maureen
; APPLICANT: DeForge,Laura
; APPLICANT: Desnoyers,Luc
; APPLICANT: Filvaroff,Ellen
; APPLICANT: Gao,Wei-Qiang
; APPLICANT: Gerritsen,Mary E.
; APPLICANT: Goddard,Audrey
; APPLICANT: Godowski,Paul J.
; APPLICANT: Gurney,Austin L.
; APPLICANT: Sherwood,Steven
; APPLICANT: Smith,Victoria
; APPLICANT: Stewart,Timothy A.
; APPLICANT: Tumas,Daniel
; APPLICANT: Watanabe,Colin K
; APPLICANT: Wood,William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3330R1C17
; CURRENT APPLICATION NUMBER: US/10/121,049
; CURRENT FILING DATE: 2002-04-12
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 361
; LENGTH: 1377
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-121-049-361

Query Match 100.0%; Score 1377; DB 9; Length 1377;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1377; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 GACTAGTTCTCTTGGAGTCTGGGAGGAGGAAAGCGGAGCCGGCAGGGAGCGAACCAGGAC 60
Db          1 GACTAGTTCTCTTGGAGTCTGGGAGGAGGAAAGCGGAGCCGGCAGGGAGCGAACCAGGAC 60

Qy         61 TGGGGTGACGGCAGGGCAGGGGGCGCCTGGCCGGGGAGAAGCGCGGGGGCTGGAGCACCA 120
Db         61 TGGGGTGACGGCAGGGCAGGGGGCGCCTGGCCGGGGAGAAGCGCGGGGGCTGGAGCACCA 120

Qy        121 CCAACTGGAGGGTCCGGAGTAGCGAGCGCCCCGAAGGAGGCCATCGGGGAGCCGGGAGGG 180
Db        121 CCAACTGGAGGGTCCGGAGTAGCGAGCGCCCCGAAGGAGGCCATCGGGGAGCCGGGAGGG 180

Qy        181 GGGACTGCGAGAGGACCCCGGCGTCCGGGCTCCCGGTGCCAGCGCTATGAGGCCACTCCT 240
Db        181 GGGACTGCGAGAGGACCCCGGCGTCCGGGCTCCCGGTGCCAGCGCTATGAGGCCACTCCT 240

Qy        241 CGTCCTGCTGCTCCTGGGCCTGGCGGCCGGCTCGCCCCCACTGGACGACAACAAGATCCC 300
Db        241 CGTCCTGCTGCTCCTGGGCCTGGCGGCCGGCTCGCCCCCACTGGACGACAACAAGATCCC 300

Qy        301 CAGCCTCTGCCCGGGGCACCCCGGCCTTCCAGGCACGCCGGGCCACCATGGCAGCCAGGG 360
Db        301 CAGCCTCTGCCCGGGGCACCCCGGCCTTCCAGGCACGCCGGGCCACCATGGCAGCCAGGG 360

Qy        361 CTTGCCGGGCCGCGATGGCCGCGACGGCCGCGACGGCGCGCCCGGGGCTCCGGGAGAGAA 420
Db        361 CTTGCCGGGCCGCGATGGCCGCGACGGCCGCGACGGCGCGCCCGGGGCTCCGGGAGAGAA 420

Qy        421 AGGCGAGGGCGGGAGGCCGGGACTGCCGGGACCTCGAGGGGACCCCGGGCCGCGAGGAGA 480
Db        421 AGGCGAGGGCGGGAGGCCGGGACTGCCGGGACCTCGAGGGGACCCCGGGCCGCGAGGAGA 480

Qy        481 GGCGGGACCCGCGGGGCCCACCGGGCCTGCCGGGGAGTGCTCGGTGCCTCCGCGATCCGC 540
Db        481 GGCGGGACCCGCGGGGCCCACCGGGCCTGCCGGGGAGTGCTCGGTGCCTCCGCGATCCGC 540

Qy        541 CTTCAGCGCCAAGCGCTCCGAGAGCCGGGTGCCTCCGCCGTCTGACGCACCCTTGCCCTT 600
Db        541 CTTCAGCGCCAAGCGCTCCGAGAGCCGGGTGCCTCCGCCGTCTGACGCACCCTTGCCCTT 600

Qy        601 CGACCGCGTGCTGGTGAACGAGCAGGGACATTACGACGCCGTCACCGGCAAGTTCACCTG 660
Db        601 CGACCGCGTGCTGGTGAACGAGCAGGGACATTACGACGCCGTCACCGGCAAGTTCACCTG 660

Qy        661 CCAGGTGCCTGGGGTCTACTACTTCGCCGTCCATGCCACCGTCTACCGGGCCAGCCTGCA 720
Db        661 CCAGGTGCCTGGGGTCTACTACTTCGCCGTCCATGCCACCGTCTACCGGGCCAGCCTGCA 720

Qy        721 GTTTGATCTGGTGAAGAATGGCGAATCCATTGCCTCTTTCTTCCAGTTTTTCGGGGGGTG 780
Db        721 GTTTGATCTGGTGAAGAATGGCGAATCCATTGCCTCTTTCTTCCAGTTTTTCGGGGGGTG 780

Qy        781 GCCCAAGCCAGCCTCGCTCTCGGGGGGGGCCATGGTGAGGCTGGAGCCTGAGGACCAAGT 840
Db        781 GCCCAAGCCAGCCTCGCTCTCGGGGGGGGCCATGGTGAGGCTGGAGCCTGAGGACCAAGT 840

Qy        841 GTGGGTGCAGGTGGGTGTGGGTGACTACATTGGCATCTATGCCAGCATCAAGACAGACAG 900
Db        841 GTGGGTGCAGGTGGGTGTGGGTGACTACATTGGCATCTATGCCAGCATCAAGACAGACAG 900

Qy        901 CACCTTCTCCGGATTTCTGGTGTACTCCGACTGGCACAGCTCCCCAGTCTTTGCTTAGTG 960
Db        901 CACCTTCTCCGGATTTCTGGTGTACTCCGACTGGCACAGCTCCCCAGTCTTTGCTTAGTG 960

Qy        961 CCCACTGCAAAGTGAGCTCATGCTCTCACTCCTAGAAGGAGGGTGTGAGGCTGACAACCA 1020
Db        961 CCCACTGCAAAGTGAGCTCATGCTCTCACTCCTAGAAGGAGGGTGTGAGGCTGACAACCA 1020

Qy       1021 GGTCATCCAGGAGGGCTGGCCCCCCCTGGAATATTGTGAATGACTAGGGAGGTGGGGTAGA 1080
Db       1021 GGTCATCCAGGAGGGCTGGCCCCCCCTGGAATATTGTGAATGACTAGGGAGGTGGGGTAGA 1080

Qy       1081 GCACTCTCCGTCCTGCTGCTGGCAAGGAATGGGAACAGTGGCTGTCTGCGATCAGGTCTG 1140
Db       1081 GCACTCTCCGTCCTGCTGCTGGCAAGGAATGGGAACAGTGGCTGTCTGCGATCAGGTCTG 1140

Qy       1141 GCAGCATGGGGCAGTGGCTGGATTTCTGCCCAAGACCAGAGGAGTGTGCTGTGCTGGCAA 1200
```

```
Db      1141 GCAGCATGGGGCAGTGGCTGGATTTCTGCCCAAGACCAGAGGAGTGTGCTGTGCTGGCAA 1200

Qy      1201 GTGTAAGTCCCCCAGTTGCTCTGGTCCAGGAGCCCACGGTGGGGTGCTCTCTTCCTGGTC 1260
Db      1201 GTGTAAGTCCCCCAGTTGCTCTGGTCCAGGAGCCCACGGTGGGGTGCTCTCTTCCTGGTC 1260

Qy      1261 CTCTGCTTCTCTGGATCCTCCCCACCCCCTCCTGCTCCTGGGGCCGGCCCTTTTCTCAGA 1320
Db      1261 CTCTGCTTCTCTGGATCCTCCCCACCCCCTCCTGCTCCTGGGGCCGGCCCTTTTCTCAGA 1320

Qy      1321 GATCACTCAATAAACCTAAGAACCCTCATAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1377
Db      1321 GATCACTCAATAAACCTAAGAACCCTCATAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1377


RESULT 9
US-10-123-904-361
; Sequence 361, Application US/10123904
; Publication No. US20030022328A1
; GENERAL INFORMATION:
;  APPLICANT: Baker,Kevin P.
;  APPLICANT:  Beresini,Maureen
;  APPLICANT:  DeForge,Laura
;  APPLICANT:  Desnoyers,Luc
;  APPLICANT:  Filvaroff,Ellen
;  APPLICANT:  Gao,Wei-Qiang
;  APPLICANT:  Gerritsen,Mary E.
;  APPLICANT:  Goddard,Audrey
;  APPLICANT:  Godowski,Paul J.
;  APPLICANT:  Gurney,Austin L.
;  APPLICANT:  Sherwood,Steven
;  APPLICANT:  Smith,Victoria
;  APPLICANT:  Stewart,Timothy A.
;  APPLICANT:  Tumas,Daniel
;  APPLICANT:  Watanabe,Colin K
;  APPLICANT:  Wood,William
;  APPLICANT:  Zhang, Zemin
;  TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;  TITLE OF INVENTION:  ACIDS ENCODING THE SAME
;  FILE REFERENCE: P3330R1C54
;  CURRENT APPLICATION NUMBER: US/10/123,904
;  CURRENT FILING DATE:  2002-04-16
;  Prior Application removed - See File Wrapper or Palm
;  NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 361
;   LENGTH: 1377
;   TYPE: DNA
;   ORGANISM: Homo Sapien
US-10-123-904-361

  Query Match           100.0%;  Score 1377;  DB 9;  Length 1377;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches 1377;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GACTAGTTCTCTTGGAGTCTGGGAGGAGGAAAGCGGAGCCGGCAGGGAGCGAACCAGGAC 60
Db          1 GACTAGTTCTCTTGGAGTCTGGGAGGAGGAAAGCGGAGCCGGCAGGGAGCGAACCAGGAC 60

Qy         61 TGGGGTGACGGCAGGGCAGGGGGCGCCTGGCCGGGGAGAAGCGCGGGGGCTGGAGCACCA 120
Db         61 TGGGGTGACGGCAGGGCAGGGGGCGCCTGGCCGGGGAGAAGCGCGGGGGCTGGAGCACCA 120

Qy        121 CCAACTGGAGGGTCCGGAGTAGCGAGCGCCCCGAAGGAGGCCATCGGGGAGCCGGGAGGG 180
Db        121 CCAACTGGAGGGTCCGGAGTAGCGAGCGCCCCGAAGGAGGCCATCGGGGAGCCGGGAGGG 180

Qy        181 GGGACTGCGAGAGGACCCCGGCGTCCGGGCTCCCGGTGCCAGCGCTATGAGGCCACTCCT 240
Db        181 GGGACTGCGAGAGGACCCCGGCGTCCGGGCTCCCGGTGCCAGCGCTATGAGGCCACTCCT 240

Qy        241 CGTCCTGCTGCTCCTGGGCCTGGCGGCCGGCTCGCCCCCACTGGACGACAACAAGATCCC 300
Db        241 CGTCCTGCTGCTCCTGGGCCTGGCGGCCGGCTCGCCCCCACTGGACGACAACAAGATCCC 300

Qy        301 CAGCCTCTGCCCCGGGGCACCCCGGCCTTCCAGGCACGCCGGGCCACCATGGCAGCCAGGG 360
Db        301 CAGCCTCTGCCCCGGGGCACCCCGGCCTTCCAGGCACGCCGGGCCACCATGGCAGCCAGGG 360

Qy        361 CTTGCCGGGCCGCGATGGCCGCGACGGCCGCGACGGCGCGCCCGGGGCTCCGGGAGAGAA 420
Db        361 CTTGCCGGGCCGCGATGGCCGCGACGGCCGCGACGGCGCGCCCGGGGCTCCGGGAGAGAA 420

Qy        421 AGGCGAGGGCGGGAGGCCGGGACTGCCGGGACCTCGAGGGGACCCCGGGCCGCGAGGAGA 480
Db        421 AGGCGAGGGCGGGAGGCCGGGACTGCCGGGACCTCGAGGGGACCCCGGGCCGCGAGGAGA 480

Qy        481 GGCGGGACCCGCGGGGCCCACCGGGCCTGCCGGGGAGTGCTCGGTGCCTCCGCGATCCGC 540
Db        481 GGCGGGACCCGCGGGGCCCACCGGGCCTGCCGGGGAGTGCTCGGTGCCTCCGCGATCCGC 540

Qy        541 CTTCAGCGCCAAGCGCTCCGAGAGCCGGGTGCCTCCGCCGTCTGACGCACCCTTGCCCTT 600
Db        541 CTTCAGCGCCAAGCGCTCCGAGAGCCGGGTGCCTCCGCCGTCTGACGCACCCTTGCCCTT 600

Qy        601 CGACCGCGTGCTGGTGAACGAGCAGGGACATTACGACGCCGTCACCGGCAAGTTCACCTG 660
Db        601 CGACCGCGTGCTGGTGAACGAGCAGGGACATTACGACGCCGTCACCGGCAAGTTCACCTG 660

Qy        661 CCAGGTGCCTGGGGTCTACTACTTCGCCGTCCATGCCACCGTCTACCGGGCCAGCCTGCA 720
Db        661 CCAGGTGCCTGGGGTCTACTACTTCGCCGTCCATGCCACCGTCTACCGGGCCAGCCTGCA 720

Qy        721 GTTTGATCTGGTGAAGAATGGCGAATCCATTGCCTCTTTCTTCCAGTTTTTCGGGGGGTG 780
Db        721 GTTTGATCTGGTGAAGAATGGCGAATCCATTGCCTCTTTCTTCCAGTTTTTCGGGGGGTG 780

Qy        781 GCCCAAGCCAGCCTCGCTCTCGGGGGGGGCCATGGTGAGGCTGGAGCCTGAGGACCAAGT 840
Db        781 GCCCAAGCCAGCCTCGCTCTCGGGGGGGGCCATGGTGAGGCTGGAGCCTGAGGACCAAGT 840

Qy        841 GTGGGTGCAGGTGGGTGTGGGTGACTACATTGGCATCTATGCCAGCATCAAGACAGACAG 900
Db        841 GTGGGTGCAGGTGGGTGTGGGTGACTACATTGGCATCTATGCCAGCATCAAGACAGACAG 900

Qy        901 CACCTTCTCCGGATTTCTGGTGTACTCCGACTGGCACAGCTCCCCAGTCTTTGCTTAGTG 960
Db        901 CACCTTCTCCGGATTTCTGGTGTACTCCGACTGGCACAGCTCCCCAGTCTTTGCTTAGTG 960

Qy        961 CCCACTGCAAAGTGAGCTCATGCTCTCACTCCTAGAAGGAGGGTGTGAGGCTGACAACCA 1020
Db        961 CCCACTGCAAAGTGAGCTCATGCTCTCACTCCTAGAAGGAGGGTGTGAGGCTGACAACCA 1020

Qy       1021 GGTCATCCAGGAGGGCTGGCCCCCCTGGAATATTGTGAATGACTAGGGAGGTGGGGTAGA 1080
Db       1021 GGTCATCCAGGAGGGCTGGCCCCCCTGGAATATTGTGAATGACTAGGGAGGTGGGGTAGA 1080

Qy       1081 GCACTCTCCGTCCTGCTGCTGGCAAGGAATGGGAACAGTGGCTGTCTGCGATCAGGTCTG 1140
Db       1081 GCACTCTCCGTCCTGCTGCTGGCAAGGAATGGGAACAGTGGCTGTCTGCGATCAGGTCTG 1140

Qy       1141 GCAGCATGGGGCAGTGGCTGGATTTCTGCCCAAGACCAGAGGAGTGTGCTGTGCTGGCAA 1200
Db       1141 GCAGCATGGGGCAGTGGCTGGATTTCTGCCCAAGACCAGAGGAGTGTGCTGTGCTGGCAA 1200

Qy       1201 GTGTAAGTCCCCCAGTTGCTCTGGTCCAGGAGCCCACGGTGGGGTGCTCTCTTCCTGGTC 1260
Db       1201 GTGTAAGTCCCCCAGTTGCTCTGGTCCAGGAGCCCACGGTGGGGTGCTCTCTTCCTGGTC 1260

Qy       1261 CTCTGCTTCTCTGGATCCTCCCCACCCCCTCCTGCTCCTGGGGCCGGCCCTTTTCTCAGA 1320
Db       1261 CTCTGCTTCTCTGGATCCTCCCCACCCCCTCCTGCTCCTGGGGCCGGCCCTTTTCTCAGA 1320

Qy       1321 GATCACTCAATAAACCTAAGAACCCTCATAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1377
Db       1321 GATCACTCAATAAACCTAAGAACCCTCATAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1377
```

```
RESULT 10
US-10-140-470-361
; Sequence 361, Application US/10140470
; Publication No. US20030022331A1
; GENERAL INFORMATION:
;  APPLICANT: Baker,Kevin P.
;  APPLICANT:  Beresini,Maureen
;  APPLICANT:  DeForge,Laura
;  APPLICANT:  Desnoyers,Luc
;  APPLICANT:  Filvaroff,Ellen
;  APPLICANT:  Gao,Wei-Qiang
;  APPLICANT:  Gerritsen,Mary E.
;  APPLICANT:  Goddard,Audrey
;  APPLICANT:  Godowski,Paul J.
;  APPLICANT:  Gurney,Austin L.
;  APPLICANT:  Sherwood,Steven
;  APPLICANT:  Smith,Victoria
;  APPLICANT:  Stewart,Timothy A.
;  APPLICANT:  Tumas,Daniel
;  APPLICANT:  Watanabe,Colin K
;  APPLICANT:  Wood,William
;  APPLICANT:  Zhang, Zemin
;  TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;  TITLE OF INVENTION:  ACIDS ENCODING THE SAME
;  FILE REFERENCE: P3330R1C160
;  CURRENT APPLICATION NUMBER: US/10/140,470
;  CURRENT FILING DATE:  2002-05-06
;  Prior Application removed - See Palm or File Wrapper
;  NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 361
;   LENGTH: 1377
;   TYPE: DNA
;   ORGANISM: Homo Sapien
US-10-140-470-361

  Query Match           100.0%;  Score 1377;  DB 9;  Length 1377;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches 1377;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GACTAGTTCTCTTGGAGTCTGGGAGGAGGAAAGCGGAGCCGGCAGGGAGCGAACCAGGAC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GACTAGTTCTCTTGGAGTCTGGGAGGAGGAAAGCGGAGCCGGCAGGGAGCGAACCAGGAC 60

Qy         61 TGGGGTGACGGCAGGGCAGGGGGCGCCTGGCCGGGGAGAAGCGCGGGGGCTGGAGCACCA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TGGGGTGACGGCAGGGCAGGGGGCGCCTGGCCGGGGAGAAGCGCGGGGGCTGGAGCACCA 120

Qy        121 CCAACTGGAGGGTCCGGAGTAGCGAGCGCCCCGAAGGAGGCCATCGGGGAGCCGGGAGGG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CCAACTGGAGGGTCCGGAGTAGCGAGCGCCCCGAAGGAGGCCATCGGGGAGCCGGGAGGG 180

Qy        181 GGGACTGCGAGAGGACCCCGGCGTCCGGGCTCCCGGTGCCAGCGCTATGAGGCCACTCCT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGGACTGCGAGAGGACCCCGGCGTCCGGGCTCCCGGTGCCAGCGCTATGAGGCCACTCCT 240

Qy        241 CGTCCTGCTGCTCCTGGGCCTGGCGGCCGGCTCGCCCCCACTGGACGACAACAAGATCCC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CGTCCTGCTGCTCCTGGGCCTGGCGGCCGGCTCGCCCCCACTGGACGACAACAAGATCCC 300

Qy        301 CAGCCTCTGCCCGGGGCACCCCGGCCTTCCAGGCACGCCGGGCCACCATGGCAGCCAGGG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CAGCCTCTGCCCGGGGCACCCCGGCCTTCCAGGCACGCCGGGCCACCATGGCAGCCAGGG 360

Qy        361 CTTGCCGGGCCGCGATGGCCGCGACGGCCGCGACGGCCGCGCCCGGGGCTCCGGGAGAGAA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CTTGCCGGGCCGCGATGGCCGCGACGGCCGCGACGGCCGCGCCCGGGGCTCCGGGAGAGAA 420

Qy        421 AGGCGAGGGCGGGAGGCCGGGACTGCCGGGACCTCGAGGGGACCCCGGGCCGCGAGGAGA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AGGCGAGGGCGGGAGGCCGGGACTGCCGGGACCTCGAGGGGACCCCGGGCCGCGAGGAGA 480

Qy        481 GGCGGGACCCGCGGGGCCCACCGGGCCTGCCGGGGAGTGCTCGGTGCCTCCGCGATCCGC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GGCGGGACCCGCGGGGCCCACCGGGCCTGCCGGGGAGTGCTCGGTGCCTCCGCGATCCGC 540

Qy        541 CTTCAGCGCCAAGCGCTCCGAGAGCCGGGTGCCTCCGCCGTCTGACGCACCCTTGCCCTT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CTTCAGCGCCAAGCGCTCCGAGAGCCGGGTGCCTCCGCCGTCTGACGCACCCTTGCCCTT 600

Qy        601 CGACCGCGTGCTGGTGAACGAGCAGGGACATTACGACGCCGTCACCGGCAAGTTCACCTG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CGACCGCGTGCTGGTGAACGAGCAGGGACATTACGACGCCGTCACCGGCAAGTTCACCTG 660

Qy        661 CCAGGTGCCTGGGGTCTACTACTTCGCCGTCCATGCCACCGTCTACCGGGCCAGCCTGCA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CCAGGTGCCTGGGGTCTACTACTTCGCCGTCCATGCCACCGTCTACCGGGCCAGCCTGCA 720

Qy        721 GTTTGATCTGGTGAAGAATGGCGAATCCATTGCCTCTTTCTTCCAGTTTTTCGGGGGGTG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GTTTGATCTGGTGAAGAATGGCGAATCCATTGCCTCTTTCTTCCAGTTTTTCGGGGGGTG 780

Qy        781 GCCCAAGCCAGCCTCGCTCTCGGGGGGGGCCATGGTGAGGCTGGAGCCTGAGGACCAAGT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GCCCAAGCCAGCCTCGCTCTCGGGGGGGGCCATGGTGAGGCTGGAGCCTGAGGACCAAGT 840

Qy        841 GTGGGTGCAGGTGGGTGTGGGTGACTACATTGGCATCTATGCCAGCATCAAGACAGACAG 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GTGGGTGCAGGTGGGTGTGGGTGACTACATTGGCATCTATGCCAGCATCAAGACAGACAG 900

Qy        901 CACCTTCTCCGGATTTCTGGTGTACTCCGACTGGCACAGCTCCCCAGTCTTTGCTTAGTG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 CACCTTCTCCGGATTTCTGGTGTACTCCGACTGGCACAGCTCCCCAGTCTTTGCTTAGTG 960

Qy        961 CCCACTGCAAAGTGAGCTCATGCTCTCACTCCTAGAAGGAGGGTGTGAGGCTGACAACCA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 CCCACTGCAAAGTGAGCTCATGCTCTCACTCCTAGAAGGAGGGTGTGAGGCTGACAACCA 1020

Qy       1021 GGTCATCCAGGAGGGCTGGCCCCCCTGGAATATTGTGAATGACTAGGGAGGTGGGGTAGA 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GGTCATCCAGGAGGGCTGGCCCCCCTGGAATATTGTGAATGACTAGGGAGGTGGGGTAGA 1080

Qy       1081 GCACTCTCCGTCCTGCTGCTGGCAAGGAATGGGAACAGTGGCTGTCTGCGATCAGGTCTG 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 GCACTCTCCGTCCTGCTGCTGGCAAGGAATGGGAACAGTGGCTGTCTGCGATCAGGTCTG 1140

Qy       1141 GCAGCATGGGGCAGTGGCTGGATTTCTGCCCAAGACCAGAGGAGTGTGCTGTGCTGGCAA 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 GCAGCATGGGGCAGTGGCTGGATTTCTGCCCAAGACCAGAGGAGTGTGCTGTGCTGGCAA 1200

Qy       1201 GTGTAAGTCCCCCAGTTGCTCTGGTCCAGGAGCCCACGGTGGGGTGCTCTCTTCCTGGTC 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 GTGTAAGTCCCCCAGTTGCTCTGGTCCAGGAGCCCACGGTGGGGTGCTCTCTTCCTGGTC 1260

Qy       1261 CTCTGCTTCTCTGGATCCTCCCCACCCCCTCCTGCTCCTGGGGCCGGCCCTTTTCTCAGA 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 CTCTGCTTCTCTGGATCCTCCCCACCCCCTCCTGCTCCTGGGGCCGGCCCTTTTCTCAGA 1320

Qy       1321 GATCACTCAATAAACCTAAGAACCCTCATAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1377
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 GATCACTCAATAAACCTAAGAACCCTCATAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1377


RESULT 11
US-10-175-746-361
; Sequence 361, Application US/10175746
; Publication No. US20030027270A1
; GENERAL INFORMATION:
;  APPLICANT: Baker,Kevin P.
;  APPLICANT:  Beresini,Maureen
;  APPLICANT:  DeForge,Laura
;  APPLICANT:  Desnoyers,Luc
;  APPLICANT:  Filvaroff,Ellen
;  APPLICANT:  Gao,Wei-Qiang
;  APPLICANT:  Gerritsen,Mary E.
```

```
;  APPLICANT:  Goddard,Audrey
;  APPLICANT:  Godowski,Paul J.
;  APPLICANT:  Gurney,Austin L.
;  APPLICANT:  Sherwood,Steven
;  APPLICANT:  Smith,Victoria
;  APPLICANT:  Stewart,Timothy A.
;  APPLICANT:  Tumas,Daniel
;  APPLICANT:  Watanabe,Colin K
;  APPLICANT:  Wood,William
;  APPLICANT:  Zhang, Zemin
;  TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;  TITLE OF INVENTION:  ACIDS ENCODING THE SAME
;  FILE REFERENCE: P3330R1C353
;  CURRENT APPLICATION NUMBER: US/10/175,746
;  CURRENT FILING DATE:  2002-06-19
;  Prior Application removed - See File Wrapper or Palm
;  NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 361
;   LENGTH: 1377
;   TYPE: DNA
;   ORGANISM: Homo Sapien
US-10-175-746-361

  Query Match           100.0%;  Score 1377;  DB 9;  Length 1377;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches 1377;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GACTAGTTCTCTTGGAGTCTGGGAGGAGGAAAGCGGAGCCGGCAGGGAGCGAACCAGGAC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GACTAGTTCTCTTGGAGTCTGGGAGGAGGAAAGCGGAGCCGGCAGGGAGCGAACCAGGAC 60

Qy         61 TGGGGTGACGGCAGGGCAGGGGGCGCCTGGCCGGGGAGAAGCGCGGGGGCTGGAGCACCA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TGGGGTGACGGCAGGGCAGGGGGCGCCTGGCCGGGGAGAAGCGCGGGGGCTGGAGCACCA 120

Qy        121 CCAACTGGAGGGTCCGGAGTAGCGAGCGCCCCGAAGGAGGCCATCGGGGAGCCGGGAGGG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CCAACTGGAGGGTCCGGAGTAGCGAGCGCCCCGAAGGAGGCCATCGGGGAGCCGGGAGGG 180

Qy        181 GGGACTGCGAGAGGACCCCGGCGTCCGGGCTCCCGGTGCCAGCGCTATGAGGCCACTCCT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGGACTGCGAGAGGACCCCGGCGTCCGGGCTCCCGGTGCCAGCGCTATGAGGCCACTCCT 240

Qy        241 CGTCCTGCTGCTCCTGGGCCTGGCGGCCGGCTCGCCCCCACTGGACGACAACAAGATCCC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CGTCCTGCTGCTCCTGGGCCTGGCGGCCGGCTCGCCCCCACTGGACGACAACAAGATCCC 300

Qy        301 CAGCCTCTGCCCGGGGCACCCCGGCCTTCCAGGCACGCCGGGCCACCATGGCAGCCAGGG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CAGCCTCTGCCCGGGGCACCCCGGCCTTCCAGGCACGCCGGGCCACCATGGCAGCCAGGG 360

Qy        361 CTTGCCGGGCCGCGATGGCCGCGACGGCCGCGACGGCGCGCCCGGGGCTCCGGGAGAGAA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CTTGCCGGGCCGCGATGGCCGCGACGGCCGCGACGGCGCGCCCGGGGCTCCGGGAGAGAA 420

Qy        421 AGGCGAGGGCGGGAGGCCGGGACTGCCGGGACCTCGAGGGGACCCCGGGCCGCGAGGAGA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AGGCGAGGGCGGGAGGCCGGGACTGCCGGGACCTCGAGGGGACCCCGGGCCGCGAGGAGA 480

Qy        481 GGCGGGACCCGCGGGGCCCACCGGGCCTGCCGGGGAGTGCTCGGTGCCTCCGCGATCCGC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GGCGGGACCCGCGGGGCCCACCGGGCCTGCCGGGGAGTGCTCGGTGCCTCCGCGATCCGC 540

Qy        541 CTTCAGCGCCAAGCGCTCCGAGAGCCGGGTGCCTCCGCCGTCTGACGCACCCTTGCCCTT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CTTCAGCGCCAAGCGCTCCGAGAGCCGGGTGCCTCCGCCGTCTGACGCACCCTTGCCCTT 600

Qy        601 CGACCGCGTGCTGGTGAACGAGCAGGGACATTACGACGCCGTCACCGGCAAGTTCACCTG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CGACCGCGTGCTGGTGAACGAGCAGGGACATTACGACGCCGTCACCGGCAAGTTCACCTG 660

Qy        661 CCAGGTGCCTGGGGTCTACTACTTCGCCGTCCATGCCACCGTCTACCGGGCCAGCCTGCA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CCAGGTGCCTGGGGTCTACTACTTCGCCGTCCATGCCACCGTCTACCGGGCCAGCCTGCA 720

Qy        721 GTTTGATCTGGTGAAGAATGGCGAATCCATTGCCTCTTTCTTCCAGTTTTTCGGGGGGTG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GTTTGATCTGGTGAAGAATGGCGAATCCATTGCCTCTTTCTTCCAGTTTTTCGGGGGGTG 780

Qy        781 GCCCAAGCCAGCCTCGCTCTCGGGGGGGGCCATGGTGAGGCTGGAGCCTGAGGACCAAGT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GCCCAAGCCAGCCTCGCTCTCGGGGGGGGCCATGGTGAGGCTGGAGCCTGAGGACCAAGT 840

Qy        841 GTGGGTGCAGGTGGGTGTGGGTGACTACATTGGCATCTATGCCAGCATCAAGACAGACAG 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GTGGGTGCAGGTGGGTGTGGGTGACTACATTGGCATCTATGCCAGCATCAAGACAGACAG 900

Qy        901 CACCTTCTCCGGATTTCTGGTGTACTCCGACTGGCACAGCTCCCCAGTCTTTGCTTAGTG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 CACCTTCTCCGGATTTCTGGTGTACTCCGACTGGCACAGCTCCCCAGTCTTTGCTTAGTG 960

Qy        961 CCCACTGCAAAGTGAGCTCATGCTCTCACTCCTAGAAGGAGGGTGTGAGGCTGACAACCA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 CCCACTGCAAAGTGAGCTCATGCTCTCACTCCTAGAAGGAGGGTGTGAGGCTGACAACCA 1020

Qy       1021 GGTCATCCAGGAGGGCTGGCCCCCCTGGAATATTGTGAATGACTAGGGAGGTGGGGTAGA 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GGTCATCCAGGAGGGCTGGCCCCCCTGGAATATTGTGAATGACTAGGGAGGTGGGGTAGA 1080

Qy       1081 GCACTCTCCGTCCTGCTGCTGGCAAGGAATGGGAACAGTGGCTGTCTGCGATCAGGTCTG 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 GCACTCTCCGTCCTGCTGCTGGCAAGGAATGGGAACAGTGGCTGTCTGCGATCAGGTCTG 1140

Qy       1141 GCAGCATGGGGCAGTGGCTGGATTTCTGCCCAAGACCAGAGGAGTGTGCTGTGCTGGCAA 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 GCAGCATGGGGCAGTGGCTGGATTTCTGCCCAAGACCAGAGGAGTGTGCTGTGCTGGCAA 1200

Qy       1201 GTGTAAGTCCCCCAGTTGCTCTGGTCCAGGAGCCCACGGTGGGGTGCTCTCTTCCTGGTC 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 GTGTAAGTCCCCCAGTTGCTCTGGTCCAGGAGCCCACGGTGGGGTGCTCTCTTCCTGGTC 1260

Qy       1261 CTCTGCTTCTCTGGATCCTCCCCACCCCCTCCTGCTCCTGGGGCCGGCCCTTTTCTCAGA 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 CTCTGCTTCTCTGGATCCTCCCCACCCCCTCCTGCTCCTGGGGCCGGCCCTTTTCTCAGA 1320

Qy       1321 GATCACTCAATAAACCTAAGAACCCTCATAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1377
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 GATCACTCAATAAACCTAAGAACCCTCATAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1377

RESULT 12
US-10-176-918-361
; Sequence 361, Application US/10176918
; Publication No. US20030027275A1
; GENERAL INFORMATION:
;  APPLICANT:  Baker,Kevin P.
;  APPLICANT:  Beresini,Maureen
;  APPLICANT:  DeForge,Laura
;  APPLICANT:  Desnoyers,Luc
;  APPLICANT:  Filvaroff,Ellen
;  APPLICANT:  Gao,Wei-Qiang
;  APPLICANT:  Gerritsen,Mary E.
;  APPLICANT:  Goddard,Audrey
;  APPLICANT:  Godowski,Paul J.
;  APPLICANT:  Gurney,Austin L.
;  APPLICANT:  Sherwood,Steven
;  APPLICANT:  Smith,Victoria
;  APPLICANT:  Stewart,Timothy A.
;  APPLICANT:  Tumas,Daniel
;  APPLICANT:  Watanabe,Colin K
;  APPLICANT:  Wood,William
;  APPLICANT:  Zhang, Zemin
;  TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;  TITLE OF INVENTION:  ACIDS ENCODING THE SAME
;  FILE REFERENCE: P3330R1C382
;  CURRENT APPLICATION NUMBER: US/10/176,918
```

```
;  CURRENT FILING DATE:  2002-06-20
;     Prior Application removed - See File Wrapper or Palm
;  NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 361
;   LENGTH: 1377
;   TYPE: DNA
;   ORGANISM: Homo Sapien
US-10-176-918-361

  Query Match           100.0%;  Score 1377;  DB 9;  Length 1377;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 1377;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GACTAGTTCTCTTGGAGTCTGGGAGGAGGAAAGCGGAGCCGGCAGGGAGCGAACCAGGAC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GACTAGTTCTCTTGGAGTCTGGGAGGAGGAAAGCGGAGCCGGCAGGGAGCGAACCAGGAC 60

Qy         61 TGGGGTGACGGCAGGGCAGGGGGCGCCTGGCCGGGGAGAAGCGCGGGGGCTGGAGCACCA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TGGGGTGACGGCAGGGCAGGGGGCGCCTGGCCGGGGAGAAGCGCGGGGGCTGGAGCACCA 120

Qy        121 CCAACTGGAGGGTCCGGAGTAGCGAGCGCCCCGAAGGAGGCCATCGGGGAGCCGGGAGGG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CCAACTGGAGGGTCCGGAGTAGCGAGCGCCCCGAAGGAGGCCATCGGGGAGCCGGGAGGG 180

Qy        181 GGGACTGCGAGAGGACCCCGGCGTCCGGGCTCCCGGTGCCAGCGCTATGAGGCCACTCCT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGGACTGCGAGAGGACCCCGGCGTCCGGGCTCCCGGTGCCAGCGCTATGAGGCCACTCCT 240

Qy        241 CGTCCTGCTGCTCCTGGGCCTGGCGGCCGGCTCGCCCCCACTGGACGACAACAAGATCCC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CGTCCTGCTGCTCCTGGGCCTGGCGGCCGGCTCGCCCCCACTGGACGACAACAAGATCCC 300

Qy        301 CAGCCTCTGCCCGGGGCACCCCGGCCTTCCAGGCACGCCGGGCCACCATGGCAGCCAGGG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CAGCCTCTGCCCGGGGCACCCCGGCCTTCCAGGCACGCCGGGCCACCATGGCAGCCAGGG 360

Qy        361 CTTGCCGGGCCGCGATGGCCGCGACGGCCGCGACGGCGCGCCCGGGGCTCCGGGAGAGAA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CTTGCCGGGCCGCGATGGCCGCGACGGCCGCGACGGCGCGCCCGGGGCTCCGGGAGAGAA 420

Qy        421 AGGCGAGGGCGGGAGGCCGGGACTGCCGGGACCTCGAGGGGACCCCGGGCCGCGAGGAGA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AGGCGAGGGCGGGAGGCCGGGACTGCCGGGACCTCGAGGGGACCCCGGGCCGCGAGGAGA 480

Qy        481 GGCGGGACCCGCGGGGCCCACCGGGCCTGCCGGGGAGTGCTCGGTGCCTCCGCGATCCGC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GGCGGGACCCGCGGGGCCCACCGGGCCTGCCGGGGAGTGCTCGGTGCCTCCGCGATCCGC 540

Qy        541 CTTCAGCGCCAAGCGCTCCGAGAGCCGGGTGCCTCCGCCGTCTGACGCACCCTTGCCCTT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CTTCAGCGCCAAGCGCTCCGAGAGCCGGGTGCCTCCGCCGTCTGACGCACCCTTGCCCTT 600

Qy        601 CGACCGCGTGCTGGTGAACGAGCAGGGACATTACGACGCCGTCACCGGCAAGTTCACCTG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CGACCGCGTGCTGGTGAACGAGCAGGGACATTACGACGCCGTCACCGGCAAGTTCACCTG 660

Qy        661 CCAGGTGCCTGGGGTCTACTACTTCGCCGTCCATGCCACCGTCTACCGGGCCAGCCTGCA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CCAGGTGCCTGGGGTCTACTACTTCGCCGTCCATGCCACCGTCTACCGGGCCAGCCTGCA 720

Qy        721 GTTTGATCTGGTGAAGAATGGCGAATCCATTGCCTCTTTCTTCCAGTTTTTCGGGGGGTG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GTTTGATCTGGTGAAGAATGGCGAATCCATTGCCTCTTTCTTCCAGTTTTTCGGGGGGTG 780

Qy        781 GCCCAAGCCAGCCTCGCTCTCGGGGGGGGCCATGGTGAGGCTGGAGCCTGAGGACCAAGT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GCCCAAGCCAGCCTCGCTCTCGGGGGGGGCCATGGTGAGGCTGGAGCCTGAGGACCAAGT 840

Qy        841 GTGGGTGCAGGTGGGTGTGGGTGACTACATTGGCATCTATGCCAGCATCAAGACAGACAG 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GTGGGTGCAGGTGGGTGTGGGTGACTACATTGGCATCTATGCCAGCATCAAGACAGACAG 900

Qy        901 CACCTTCTCCGGATTTCTGGTGTACTCCGACTGGCACAGCTCCCCAGTCTTTGCTTAGTG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 CACCTTCTCCGGATTTCTGGTGTACTCCGACTGGCACAGCTCCCCAGTCTTTGCTTAGTG 960

Qy        961 CCCACTGCAAAGTGAGCTCATGCTCTCACTCCTAGAAGGAGGGTGTGAGGCTGACAACCA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 CCCACTGCAAAGTGAGCTCATGCTCTCACTCCTAGAAGGAGGGTGTGAGGCTGACAACCA 1020

Qy       1021 GGTCATCCAGGAGGGCTGGCCCCCCTGGAATATTGTGAATGACTAGGGAGGTGGGGTAGA 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GGTCATCCAGGAGGGCTGGCCCCCCTGGAATATTGTGAATGACTAGGGAGGTGGGGTAGA 1080

Qy       1081 GCACTCTCCGTCCTGCTGCTGGCAAGGAATGGGAACAGTGGCTGTCTGCGATCAGGTCTG 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 GCACTCTCCGTCCTGCTGCTGGCAAGGAATGGGAACAGTGGCTGTCTGCGATCAGGTCTG 1140

Qy       1141 GCAGCATGGGGCAGTGGCTGGATTTCTGCCCAAGACCAGAGGAGTGTGCTGTGCTGGCAA 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 GCAGCATGGGGCAGTGGCTGGATTTCTGCCCAAGACCAGAGGAGTGTGCTGTGCTGGCAA 1200

Qy       1201 GTGTAAGTCCCCCAGTTGCTCTGGTCCAGGAGCCCACGGTGGGGTGCTCTCTTCCTGGTC 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 GTGTAAGTCCCCCAGTTGCTCTGGTCCAGGAGCCCACGGTGGGGTGCTCTCTTCCTGGTC 1260

Qy       1261 CTCTGCTTCTCTGGATCCTCCCCACCCCCTCCTGCTCCTGGGGCCGGCCCTTTTCTCAGA 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 CTCTGCTTCTCTGGATCCTCCCCACCCCCTCCTGCTCCTGGGGCCGGCCCTTTTCTCAGA 1320

Qy       1321 GATCACTCAATAAACCTAAGAACCCTCATAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1377
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 GATCACTCAATAAACCTAAGAACCCTCATAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1377


RESULT 13
US-10-176-921-361
; Sequence 361, Application US/10176921
; Publication No. US20030027276A1
; GENERAL INFORMATION:
;  APPLICANT: Baker,Kevin P.
;  APPLICANT:  Beresini,Maureen
;  APPLICANT:  DeForge,Laura
;  APPLICANT:  Desnoyers,Luc
;  APPLICANT:  Filvaroff,Ellen
;  APPLICANT:  Gao,Wei-Qiang
;  APPLICANT:  Gerritsen,Mary E.
;  APPLICANT:  Goddard,Audrey
;  APPLICANT:  Godowski,Paul J.
;  APPLICANT:  Gurney,Austin L.
;  APPLICANT:  Sherwood,Steven
;  APPLICANT:  Smith,Victoria
;  APPLICANT:  Stewart,Timothy A.
;  APPLICANT:  Tumas,Daniel
;  APPLICANT:  Watanabe,Colin K
;  APPLICANT:  Wood,William
;  APPLICANT:  Zhang, Zemin
;  TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;  TITLE OF INVENTION:  ACIDS ENCODING THE SAME
;  FILE REFERENCE: P3330R1C288
;  CURRENT APPLICATION NUMBER: US/10/176,921
;  CURRENT FILING DATE:  2002-06-20
;  Prior Application removed - See File Wrapper or Palm
;  NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 361
;   LENGTH: 1377
;   TYPE: DNA
;   ORGANISM: Homo Sapien
US-10-176-921-361

  Query Match           100.0%;  Score 1377;  DB 9;  Length 1377;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 1377;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GACTAGTTCTCTTGGAGTCTGGGAGGAGGAAAGCGGAGCCGGCAGGGAGCGAACCAGGAC 60
```

```
Db          1 GACTAGTTCTCTTGGAGTCTGGGAGGAGGAAAGCGGAGCCGGCAGGGAGCGAACCAGGAC 60
Qy         61 TGGGGTGACGGCAGGGCAGGGGGCGCCTGGCCGGGGAGAAGCGCGGGGGCTGGAGCACCA 120
Db         61 TGGGGTGACGGCAGGGCAGGGGGCGCCTGGCCGGGGAGAAGCGCGGGGGCTGGAGCACCA 120
Qy        121 CCAACTGGAGGGTCCGGAGTAGCGAGCGCCCCGAAGGAGGCCATCGGGGAGCCGGGAGGG 180
Db        121 CCAACTGGAGGGTCCGGAGTAGCGAGCGCCCCGAAGGAGGCCATCGGGGAGCCGGGAGGG 180
Qy        181 GGGACTGCGAGAGGACCCCGGCGTCCGGGCTCCCGGTGCCAGCGCTATGAGGCCACTCCT 240
Db        181 GGGACTGCGAGAGGACCCCGGCGTCCGGGCTCCCGGTGCCAGCGCTATGAGGCCACTCCT 240
Qy        241 CGTCCTGCTGCTCCTGGGCCTGGCGGCCGGCTCGCCCCCACTGGACGACAACAAGATCCC 300
Db        241 CGTCCTGCTGCTCCTGGGCCTGGCGGCCGGCTCGCCCCCACTGGACGACAACAAGATCCC 300
Qy        301 CAGCCTCTGCCCGGGGCACCCCGGCCTTCCAGGCACGCCGGGCCACCATGGCAGCCAGGG 360
Db        301 CAGCCTCTGCCCGGGGCACCCCGGCCTTCCAGGCACGCCGGGCCACCATGGCAGCCAGGG 360
Qy        361 CTTGCCGGGCCGCGATGGCCGCGACGGCCGCGACGGCGCGCCCGGGGCTCCGGGAGAGAA 420
Db        361 CTTGCCGGGCCGCGATGGCCGCGACGGCCGCGACGGCGCGCCCGGGGCTCCGGGAGAGAA 420
Qy        421 AGGCGAGGGCGGGAGGCCGGGACTGCCGGGACCTCGAGGGGACCCCGGGCCGCGAGGAGA 480
Db        421 AGGCGAGGGCGGGAGGCCGGGACTGCCGGGACCTCGAGGGGACCCCGGGCCGCGAGGAGA 480
Qy        481 GGCGGGACCCGCGGGGGCCCACCGGGCCTGCCGGGGAGTGCTCGGTGCCTCCGCGATCCGC 540
Db        481 GGCGGGACCCGCGGGGGCCCACCGGGCCTGCCGGGGAGTGCTCGGTGCCTCCGCGATCCGC 540
Qy        541 CTTCAGCGCCAAGCGCTCCGAGAGCCGGGTGCCTCCGCCGTCTGACGCACCCTTGCCCTT 600
Db        541 CTTCAGCGCCAAGCGCTCCGAGAGCCGGGTGCCTCCGCCGTCTGACGCACCCTTGCCCTT 600
Qy        601 CGACCGCGTGCTGGTGAACGAGCAGGGACATTACGACGCCGTCACCGGCAAGTTCACCTG 660
Db        601 CGACCGCGTGCTGGTGAACGAGCAGGGACATTACGACGCCGTCACCGGCAAGTTCACCTG 660
Qy        661 CCAGGTGCCTGGGGTCTACTACTTCGCCGTCCATGCCACCGTCTACCGGGCCAGCCTGCA 720
Db        661 CCAGGTGCCTGGGGTCTACTACTTCGCCGTCCATGCCACCGTCTACCGGGCCAGCCTGCA 720
Qy        721 GTTTGATCTGGTGAAGAATGGCGAATCCATTGCCTCTTTCTTCCAGTTTTTCGGGGGGTG 780
Db        721 GTTTGATCTGGTGAAGAATGGCGAATCCATTGCCTCTTTCTTCCAGTTTTTCGGGGGGTG 780
Qy        781 GCCCAAGCCAGCCTCGCTCTCGGGGGGGGCCATGGTGAGGCTGGAGCCTGAGGACCAAGT 840
Db        781 GCCCAAGCCAGCCTCGCTCTCGGGGGGGGCCATGGTGAGGCTGGAGCCTGAGGACCAAGT 840
Qy        841 GTGGGTGCAGGTGGGTGTGGGTGACTACATTGGCATCTATGCCAGCATCAAGACAGACAG 900
Db        841 GTGGGTGCAGGTGGGTGTGGGTGACTACATTGGCATCTATGCCAGCATCAAGACAGACAG 900
Qy        901 CACCTTCTCCGGATTTCTGGTGTACTCCGACTGGCACAGCTCCCCAGTCTTTGCTTAGTG 960
Db        901 CACCTTCTCCGGATTTCTGGTGTACTCCGACTGGCACAGCTCCCCAGTCTTTGCTTAGTG 960
Qy        961 CCCACTGCAAAGTGAGCTCATGCTCTCACTCCTAGAAGGAGGGTGTGAGGCTGACAACCA 1020
Db        961 CCCACTGCAAAGTGAGCTCATGCTCTCACTCCTAGAAGGAGGGTGTGAGGCTGACAACCA 1020
Qy       1021 GGTCATCCAGGAGGGCTGGCCCCCCTGGAATATTGTGAATGACTAGGGAGGTGGGGTAGA 1080
Db       1021 GGTCATCCAGGAGGGCTGGCCCCCCTGGAATATTGTGAATGACTAGGGAGGTGGGGTAGA 1080
Qy       1081 GCACTCTCCGTCCTGCTGCTGGCAAGGAATGGGAACAGTGGCTGTCTGCGATCAGGTCTG 1140
Db       1081 GCACTCTCCGTCCTGCTGCTGGCAAGGAATGGGAACAGTGGCTGTCTGCGATCAGGTCTG 1140
Qy       1141 GCAGCATGGGGCAGTGGCTGGATTTCTGCCCAAGACCAGAGGAGTGTGCTGTGCTGGCAA 1200
Db       1141 GCAGCATGGGGCAGTGGCTGGATTTCTGCCCAAGACCAGAGGAGTGTGCTGTGCTGGCAA 1200
Qy       1201 GTGTAAGTCCCCCAGTTGCTCTGGTCCAGGAGCCCACGGTGGGGTGCTCTCTTCCTGGTC 1260
Db       1201 GTGTAAGTCCCCCAGTTGCTCTGGTCCAGGAGCCCACGGTGGGGTGCTCTCTTCCTGGTC 1260
Qy       1261 CTCTGCTTCTCTGGATCCTCCCCACCCCCTCCTGCTCCTGGGGCCGGCCCTTTTCTCAGA 1320
Db       1261 CTCTGCTTCTCTGGATCCTCCCCACCCCCTCCTGCTCCTGGGGCCGGCCCTTTTCTCAGA 1320
Qy       1321 GATCACTCAATAAACCTAAGAACCCTCATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1377
Db       1321 GATCACTCAATAAACCTAAGAACCCTCATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1377
```

```
RESULT 14
US-10-137-865-361
; Sequence 361, Application US/10137865
; Publication No. US20030032155A1
; GENERAL INFORMATION:
;  APPLICANT: Baker,Kevin P.
;  APPLICANT:  Beresini,Maureen
;  APPLICANT:  DeForge,Laura
;  APPLICANT:  Desnoyers,Luc
;  APPLICANT:  Filvaroff,Ellen
;  APPLICANT:  Gao,Wei-Qiang
;  APPLICANT:  Gerritsen,Mary E.
;  APPLICANT:  Goddard,Audrey
;  APPLICANT:  Godowski,Paul J.
;  APPLICANT:  Gurney,Austin L.
;  APPLICANT:  Sherwood,Steven
;  APPLICANT:  Smith,Victoria
;  APPLICANT:  Stewart,Timothy A.
;  APPLICANT:  Tumas,Daniel
;  APPLICANT:  Watanabe,Colin K
;  APPLICANT:  Wood,William
;  APPLICANT:  Zhang, Zemin
;  TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;  TITLE OF INVENTION:  ACIDS ENCODING THE SAME
;  FILE REFERENCE: P3330R1C154
;  CURRENT APPLICATION NUMBER: US/10/137,865
;  CURRENT FILING DATE:  2002-05-03
;  Prior Application removed - See Palm or File Wrapper
;  NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 361
;   LENGTH: 1377
;   TYPE: DNA
;   ORGANISM: Homo Sapien
US-10-137-865-361

  Query Match           100.0%;  Score 1377;  DB 9;  Length 1377;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches 1377;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GACTAGTTCTCTTGGAGTCTGGGAGGAGGAAAGCGGAGCCGGCAGGGAGCGAACCAGGAC 60
Db          1 GACTAGTTCTCTTGGAGTCTGGGAGGAGGAAAGCGGAGCCGGCAGGGAGCGAACCAGGAC 60
Qy         61 TGGGGTGACGGCAGGGCAGGGGGCGCCTGGCCGGGGAGAAGCGCGGGGGCTGGAGCACCA 120
Db         61 TGGGGTGACGGCAGGGCAGGGGGCGCCTGGCCGGGGAGAAGCGCGGGGGCTGGAGCACCA 120
Qy        121 CCAACTGGAGGGTCCGGAGTAGCGAGCGCCCCGAAGGAGGCCATCGGGGAGCCGGGAGGG 180
Db        121 CCAACTGGAGGGTCCGGAGTAGCGAGCGCCCCGAAGGAGGCCATCGGGGAGCCGGGAGGG 180
Qy        181 GGGACTGCGAGAGGACCCCGGCGTCCGGGCTCCCGGTGCCAGCGCTATGAGGCCACTCCT 240
Db        181 GGGACTGCGAGAGGACCCCGGCGTCCGGGCTCCCGGTGCCAGCGCTATGAGGCCACTCCT 240
```

```
Qy       241 CGTCCTGCTGCTCCTGGGCCTGGCGGCCGGCTCGCCCCCACTGGACGACAACAAGATCCC 300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       241 CGTCCTGCTGCTCCTGGGCCTGGCGGCCGGCTCGCCCCCACTGGACGACAACAAGATCCC 300

Qy       301 CAGCCTCTGCCCGGGGCACCCCGGCCTTCCAGGCACGCCGGGCCACCATGGCAGCCAGGG 360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       301 CAGCCTCTGCCCGGGGCACCCCGGCCTTCCAGGCACGCCGGGCCACCATGGCAGCCAGGG 360

Qy       361 CTTGCCGGGCCGCGATGGCCGCGACGGCCGCGACGGCGCGCCCGGGGCTCCGGGAGAGAA 420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       361 CTTGCCGGGCCGCGATGGCCGCGACGGCCGCGACGGCGCGCCCGGGGCTCCGGGAGAGAA 420

Qy       421 AGGCGAGGGCGGGAGGCCGGGACTGCCGGGACCTCGAGGGGACCCCGGGCCGCGAGGAGA 480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       421 AGGCGAGGGCGGGAGGCCGGGACTGCCGGGACCTCGAGGGGACCCCGGGCCGCGAGGAGA 480

Qy       481 GGCGGGACCCGCGGGGCCCACCGGGCCTGCCGGGGAGTGCTCGGTGCCTCCGCGATCCGC 540
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       481 GGCGGGACCCGCGGGGCCCACCGGGCCTGCCGGGGAGTGCTCGGTGCCTCCGCGATCCGC 540

Qy       541 CTTCAGCGCCAAGCGCTCCGAGAGCCGGGTGCCTCCGCCGTCTGACGCACCCTTGCCCTT 600
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       541 CTTCAGCGCCAAGCGCTCCGAGAGCCGGGTGCCTCCGCCGTCTGACGCACCCTTGCCCTT 600

Qy       601 CGACCGCGTGCTGGTGAACGAGCAGGGACATTACGACGCCGTCACCGGCAAGTTCACCTG 660
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       601 CGACCGCGTGCTGGTGAACGAGCAGGGACATTACGACGCCGTCACCGGCAAGTTCACCTG 660

Qy       661 CCAGGTGCCTGGGGTCTACTACTTCGCCGTCCATGCCACCGTCTACCGGGCCAGCCTGCA 720
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       661 CCAGGTGCCTGGGGTCTACTACTTCGCCGTCCATGCCACCGTCTACCGGGCCAGCCTGCA 720

Qy       721 GTTTGATCTGGTGAAGAATGGCGAATCCATTGCCTCTTTCTTCCAGTTTTTCGGGGGGTG 780
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       721 GTTTGATCTGGTGAAGAATGGCGAATCCATTGCCTCTTTCTTCCAGTTTTTCGGGGGGTG 780

Qy       781 GCCCAAGCCAGCCTCGCTCTCGGGGGGGGCCATGGTGAGGCTGGAGCCTGAGGACCAAGT 840
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       781 GCCCAAGCCAGCCTCGCTCTCGGGGGGGGCCATGGTGAGGCTGGAGCCTGAGGACCAAGT 840

Qy       841 GTGGGTGCAGGTGGGTGTGGGTGACTACATTGGCATCTATGCCAGCATCAAGACAGACAG 900
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       841 GTGGGTGCAGGTGGGTGTGGGTGACTACATTGGCATCTATGCCAGCATCAAGACAGACAG 900

Qy       901 CACCTTCTCCGGATTTCTGGTGTACTCCGACTGGCACAGCTCCCCAGTCTTTGCTTAGTG 960
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       901 CACCTTCTCCGGATTTCTGGTGTACTCCGACTGGCACAGCTCCCCAGTCTTTGCTTAGTG 960

Qy       961 CCCACTGCAAAGTGAGCTCATGCTCTCACTCCTAGAAGGAGGGTGTGAGGCTGACAACCA 1020
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       961 CCCACTGCAAAGTGAGCTCATGCTCTCACTCCTAGAAGGAGGGTGTGAGGCTGACAACCA 1020

Qy      1021 GGTCATCCAGGAGGGCTGGCCCCCCTGGAATATTGTGAATGACTAGGGAGGTGGGGTAGA 1080
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1021 GGTCATCCAGGAGGGCTGGCCCCCCTGGAATATTGTGAATGACTAGGGAGGTGGGGTAGA 1080

Qy      1081 GCACTCTCCGTCCTGCTGCTGGCAAGGAATGGGAACAGTGGCTGTCTGCGATCAGGTCTG 1140
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1081 GCACTCTCCGTCCTGCTGCTGGCAAGGAATGGGAACAGTGGCTGTCTGCGATCAGGTCTG 1140

Qy      1141 GCAGCATGGGGCAGTGGCTGGATTTCTGCCCAAGACCAGAGGAGTGTGCTGTGCTGGCAA 1200
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1141 GCAGCATGGGGCAGTGGCTGGATTTCTGCCCAAGACCAGAGGAGTGTGCTGTGCTGGCAA 1200

Qy      1201 GTGTAAGTCCCCCAGTTGCTCTGGTCCAGGAGCCCACGGTGGGGTGCTCTCTTCCTGGTC 1260
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1201 GTGTAAGTCCCCCAGTTGCTCTGGTCCAGGAGCCCACGGTGGGGTGCTCTCTTCCTGGTC 1260

Qy      1261 CTCTGCTTCTCTGGATCCTCCCCACCCCCTCCTGCTCCTGGGGCCGGCCCTTTTCTCAGA 1320
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1261 CTCTGCTTCTCTGGATCCTCCCCACCCCCTCCTGCTCCTGGGGCCGGCCCTTTTCTCAGA 1320

Qy      1321 GATCACTCAATAAACCTAAGAACCCTCATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1377
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1321 GATCACTCAATAAACCTAAGAACCCTCATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1377
```

RESULT 15
US-10-140-474-361
; Sequence 361, Application US/10140474
; Publication No. US20030032156A1
; GENERAL INFORMATION:
; APPLICANT: Baker,Kevin P.
; APPLICANT: Beresini,Maureen
; APPLICANT: DeForge,Laura
; APPLICANT: Desnoyers,Luc
; APPLICANT: Filvaroff,Ellen
; APPLICANT: Gao,Wei-Qiang
; APPLICANT: Gerritsen,Mary E.
; APPLICANT: Goddard,Audrey
; APPLICANT: Godowski,Paul J.
; APPLICANT: Gurney,Austin L.
; APPLICANT: Sherwood,Steven
; APPLICANT: Smith,Victoria
; APPLICANT: Stewart,Timothy A.
; APPLICANT: Tumas,Daniel
; APPLICANT: Watanabe,Colin K
; APPLICANT: Wood,William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3330R1C162
; CURRENT APPLICATION NUMBER: US/10/140,474
; CURRENT FILING DATE: 2002-05-06
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 361
; LENGTH: 1377
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-140-474-361

Query Match 100.0%; Score 1377; DB 9; Length 1377;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1377; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 GACTAGTTCTCTTGGAGTCTGGGAGGAGGAAAGCGGAGCCGGCAGGGAGCGAACCAGGAC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GACTAGTTCTCTTGGAGTCTGGGAGGAGGAAAGCGGAGCCGGCAGGGAGCGAACCAGGAC 60

Qy         61 TGGGGTGACGGCAGGGCAGGGGGCGCCTGGCCGGGGAGAAGCGCGGGGGCTGGAGCACCA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TGGGGTGACGGCAGGGCAGGGGGCGCCTGGCCGGGGAGAAGCGCGGGGGCTGGAGCACCA 120

Qy        121 CCAACTGGAGGGTCCGGAGTAGCGAGCGCCCCGAAGGAGGCCATCGGGGAGCCGGGAGGG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CCAACTGGAGGGTCCGGAGTAGCGAGCGCCCCGAAGGAGGCCATCGGGGAGCCGGGAGGG 180

Qy        181 GGGACTGCGAGAGGACCCCGGCGTCCGGGCTCCCGGTGCCAGCGCTATGAGGCCACTCCT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGGACTGCGAGAGGACCCCGGCGTCCGGGCTCCCGGTGCCAGCGCTATGAGGCCACTCCT 240

Qy        241 CGTCCTGCTGCTCCTGGGCCTGGCGGCCGGCTCGCCCCCACTGGACGACAACAAGATCCC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CGTCCTGCTGCTCCTGGGCCTGGCGGCCGGCTCGCCCCCACTGGACGACAACAAGATCCC 300

Qy        301 CAGCCTCTGCCCGGGGCACCCCGGCCTTCCAGGCACGCCGGGCCACCATGGCAGCCAGGG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CAGCCTCTGCCCGGGGCACCCCGGCCTTCCAGGCACGCCGGGCCACCATGGCAGCCAGGG 360

Qy        361 CTTGCCGGGCCGCGATGGCCGCGACGGCCGCGACGGCGCGCCCGGGGCTCCGGGAGAGAA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CTTGCCGGGCCGCGATGGCCGCGACGGCCGCGACGGCGCGCCCGGGGCTCCGGGAGAGAA 420

Qy        421 AGGCGAGGGCGGGAGGCCGGGACTGCCGGGACCTCGAGGGGACCCCGGGCCGCGAGGAGA 480
```

```
Db        421 AGGCGAGGGCGGGAGGCCGGGACTGCCGGGACCTCGAGGGGACCCCGGGCCGCGAGGAGA 480

Qy        481 GGCGGGACCCGCGGGGCCCACCGGGCCTGCCGGGGAGTGCTCGGTGCCTCCGCGATCCGC 540

Db        481 GGCGGGACCCGCGGGGCCCACCGGGCCTGCCGGGGAGTGCTCGGTGCCTCCGCGATCCGC 540

Qy        541 CTTCAGCGCCAAGCGCTCCGAGAGCCGGGTGCCTCCGCCGTCTGACGCACCCTTGCCCTT 600

Db        541 CTTCAGCGCCAAGCGCTCCGAGAGCCGGGTGCCTCCGCCGTCTGACGCACCCTTGCCCTT 600

Qy        601 CGACCGCGTGCTGGTGAACGAGCAGGGACATTACGACGCCGTCACCGGCAAGTTCACCTG 660

Db        601 CGACCGCGTGCTGGTGAACGAGCAGGGACATTACGACGCCGTCACCGGCAAGTTCACCTG 660

Qy        661 CCAGGTGCCTGGGGTCTACTACTTCGCCGTCCATGCCACCGTCTACCGGGCCAGCCTGCA 720

Db        661 CCAGGTGCCTGGGGTCTACTACTTCGCCGTCCATGCCACCGTCTACCGGGCCAGCCTGCA 720

Qy        721 GTTTGATCTGGTGAAGAATGGCGAATCCATTGCCTCTTTCTTCCAGTTTTTCGGGGGGTG 780

Db        721 GTTTGATCTGGTGAAGAATGGCGAATCCATTGCCTCTTTCTTCCAGTTTTTCGGGGGGTG 780

Qy        781 GCCCAAGCCAGCCTCGCTCTCGGGGGGGGCCATGGTGAGGCTGGAGCCTGAGGACCAAGT 840

Db        781 GCCCAAGCCAGCCTCGCTCTCGGGGGGGGCCATGGTGAGGCTGGAGCCTGAGGACCAAGT 840

Qy        841 GTGGGTGCAGGTGGGTGTGGGTGACTACATTGGCATCTATGCCAGCATCAAGACAGACAG 900

Db        841 GTGGGTGCAGGTGGGTGTGGGTGACTACATTGGCATCTATGCCAGCATCAAGACAGACAG 900

Qy        901 CACCTTCTCCGGATTTCTGGTGTACTCCGACTGGCACAGCTCCCCAGTCTTTGCTTAGTG 960

Db        901 CACCTTCTCCGGATTTCTGGTGTACTCCGACTGGCACAGCTCCCCAGTCTTTGCTTAGTG 960

Qy        961 CCCACTGCAAAGTGAGCTCATGCTCTCACTCCTAGAAGGAGGGTGTGAGGCTGACAACCA 1020

Db        961 CCCACTGCAAAGTGAGCTCATGCTCTCACTCCTAGAAGGAGGGTGTGAGGCTGACAACCA 1020

Qy       1021 GGTCATCCAGGAGGGCTGGCCCCCCCTGGAATATTGTGAATGACTAGGGAGGTGGGGTAGA 1080

Db       1021 GGTCATCCAGGAGGGCTGGCCCCCCCTGGAATATTGTGAATGACTAGGGAGGTGGGGTAGA 1080

Qy       1081 GCACTCTCCGTCCTGCTGCTGGCAAGGAATGGGAACAGTGGCTGTCTGCGATCAGGTCTG 1140

Db       1081 GCACTCTCCGTCCTGCTGCTGGCAAGGAATGGGAACAGTGGCTGTCTGCGATCAGGTCTG 1140

Qy       1141 GCAGCATGGGGCAGTGGCTGGATTTCTGCCCAAGACCAGAGGAGTGTGCTGTGCTGGCAA 1200

Db       1141 GCAGCATGGGGCAGTGGCTGGATTTCTGCCCAAGACCAGAGGAGTGTGCTGTGCTGGCAA 1200

Qy       1201 GTGTAAGTCCCCCAGTTGCTCTGGTCCAGGAGCCCACGGTGGGGTGCTCTCTTCCTGGTC 1260

Db       1201 GTGTAAGTCCCCCAGTTGCTCTGGTCCAGGAGCCCACGGTGGGGTGCTCTCTTCCTGGTC 1260

Qy       1261 CTCTGCTTCTCTGGATCCTCCCCACCCCCTCCTGCTCCTGGGGCCGGCCCTTTTCTCAGA 1320

Db       1261 CTCTGCTTCTCTGGATCCTCCCCACCCCCTCCTGCTCCTGGGGCCGGCCCTTTTCTCAGA 1320

Qy       1321 GATCACTCAATAAACCTAAGAACCCTCATAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1377

Db       1321 GATCACTCAATAAACCTAAGAACCCTCATAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1377

Search completed: June 22, 2003, 06:27:28
Job time : 229 secs
```